型番:SR-MOD100(Rev.B)　　REV1.70

小型・お手軽
簡単!UART音声認識組込モジュール

取扱説明書
お使いになる前にこの説明書をよくお読みの上正しくお使いください。

(C)2015 マイクロテクニカ

※本マニュアルは、製品本体基板裏面に"EasyVR"と記載のある製品用です。"VRbot"と記載の製品は別のマニュアルとなります。

製品の概要

簡単!UART音声認識組込モジュール(型式:SR-MOD100)は、45mm×24mmの小さな基板に音声認識に必要な機能を凝縮した組込用の音声認識モジュールです。最初から用意されている基本フレーズ25語の他に、ユーザーが任意のフレーズを32フレーズ登録(※1)させ、音声認識をさせることができます。(認識フレーズは、フレーズを学習させた人の声しか認識できません。)

日本語はもちろん、英語・イタリア語・ドイツ語・フランス語・スペイン語の6カ国語に対応しています。登録させたフレーズ情報はすべて本体の不揮発性メモリーに記憶されます。

制御はすべて簡単な非同期式シリアル通信(UART)で行うことができます。音声認識の開始コマンドを送信後、マイクに入力された音声を解析し、以前に学習させた音声と認識した場合には、シリアルコマンドで対応するインデックス番号を返します。ユーザーは、このインデックス番号によって、どの音声が入力されたかを認識することができます。

SR-MOD100の操作はすべて簡単なシリアルコマンドだけで構成されています。ユーザーは音声認識に関する特別な知識は必要ありません。UARTのロジック電圧レベルは+5Vと+3.3V両方に対応していますので、PICマイコンなどのワンチップマイコンと接続して簡単に音声認識アプリケーションを作成できます。

認識させる音声の学習や管理は、パソコン上から簡単に行える専用のユーティリティソフトが付属しています。別売のパソコン接続キットを使うことで、パソコンのRS232Cポートに接続して簡単にパソコンから制御することができます。

その他、コマンドで制御できる3ビットのI/Oピンも搭載しています。

※音声認識には、発声した声質やイントネーション、その他背景の雑音等の諸条件によって正しく音声を認識できない場合があります。これは製品の性質上やむを得ない事象であり、本製品の音声認識精度は保証されるものではありません。

※本製品は伊RoboTech srl社製の製品です。

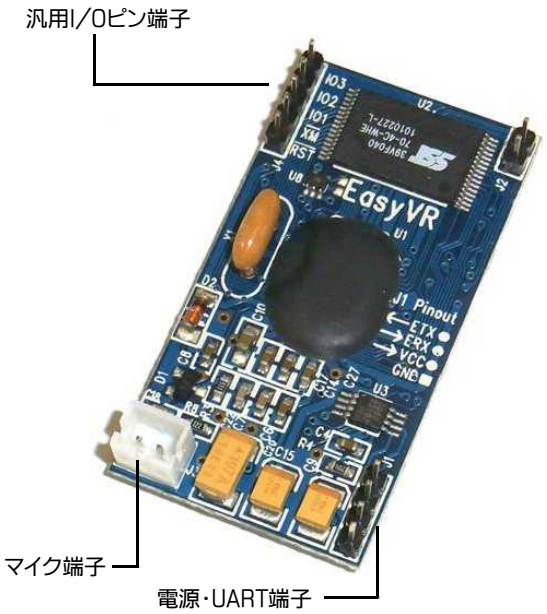
※1:32フレーズとは、話者特定パスワードフレーズ(話者識別認識)の最大登録可能数5人分を除いた場合のフレーズ数です。パスワードフレーズを5つ登録した場合、話者特定フレーズは最大27フレーズまでとなります。(話者特定パスワードフレーズを使用しない場合には、話者特定フレーズを最大32フレーズ登録できます。)

パッケージの内容

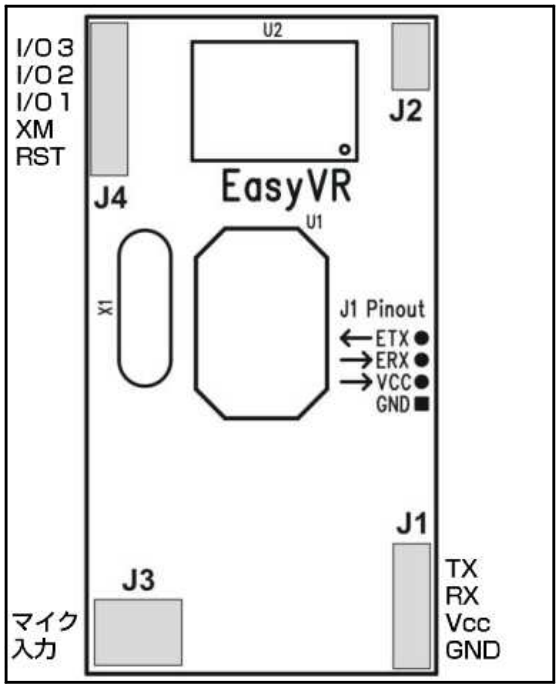
■同梱物
- ・SR-MOD100本体
- ・2ピンハーネス　2本(電源用、通信線用)
- ・コンデンサマイクケーブル

■インターネットからダウンロードして頂くもの
- ・マニュアル(本書)
- ・ST-MOD100用ユーティリティソフト

ピン配置及びピンの詳細



※使用しないピン端子には何も接続しないでください。電圧が印加されたり、GNDに接続すると本体が破損することがあります。



■電源・UART端子(J1)

基板の下側から順番に1ピン～4ピンとなっています。ピンの詳細は下記の通りです。

ピン	詳　細
4	UART－送信データ端子(TX) TTLレベル
3	UART－受信データ端子(RX) TTLレベル
2	電源端子　+3.3V～+5.0V
1	電源端子　GND

■コンデンサマイク端子(J3)

基板の端側がGND側になります。コンデンサマイクを取り付ける場合には、極性にご注意ください。ピンのピッチは2mmピッチです。

■汎用I/Oピン端子(J4)

基板の下側から順番に1ピン～4ピンとなっています。ピンの詳細は下記の通りです。

ピン	詳　細
5	汎用I/Oピン3　(※注意!　最大3.0Vまで)
4	汎用I/Oピン2　(※注意!　最大3.0Vまで)
3	汎用I/Oピン1　(※注意!　最大3.0Vまで)
2	使用しません　(ブートセレクト)　何も接続しないでください
1	使用しません　(UARTリセット)　何も接続しないでください

※UARTピンは最大Vcc電圧まで印加できますが、汎用I/Oピンは3.0Vが最大印加可能電圧です。これを上回る電圧を印加すると、本体は破損します。十分ご注意ください。

※J2ピンは使用しません。何も接続しないでください。

各ピンの電気的特性

■推奨動作条件

表示	詳細	Min	Typ	Max	単位
Vcc	電源電圧	3.3	5.0	5.5	V
Ta	周囲温度	0	25	70	℃
RX	シリアル通信　受信データ	0	－	Vcc	V
TX	シリアル通信　送信データ	0	－	Vcc	V

■J4端子(汎用I/Oピン)電気的特性

表示	詳細		Min	Typ	Max	単位
VIH	入力H電圧		2.4	3.0	3.3	V
VIL	入力L電圧		−0.1	0.0	0.75	V
IIL	入力漏れ電流 (0V<Vio<3V Hi-Z input)			<1	10	μA
Rpu	ﾌﾟﾙｱｯﾌﾟ抵抗	Strong		10		kΩ
		Weak		200		kΩ
VOH	出力H電圧 (IOH=−5mA)		2.4			V
VOL	出力L電圧 (IOL=8mA)				0.6	V

■電源仕様

表示	詳細	Min	Typ	Max	単位
Isp	スリープ時電流	<1			mA
Iop	動作時電流	12			mA

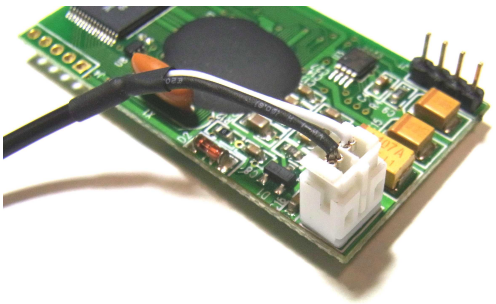
SR-MOD100を準備しましょう

SR−MOD100を使用できるよう準備をしましょう。SR−MOD100に接続する非同期式シリアル通信(以下、UARTと記載)のホスト機器は、PICマイコンなどのマイコンが使用できますが、最初に音声認識用のフレーズを本体に記憶、学習させるために、パソコンとの接続が必要です。

SR−MOD100のUART端子のロジック電圧レベルは0VとVccに印加した電圧レベルとなります。本体の電源電圧は+3.3V又は+5.0Vどちらでも動作しますので、電源電圧が+3.3Vの時は+3.3V系のロジック電圧レベル、+5Vの時は+5V系のロジック電圧レベルでご使用ください。

①コンデンサマイクを接続しましょう

付属のコンデンサマイクをSR−MOD100のコンデンサマイク端子に接続します。端子には切り込みがありますので、切り込みに合わせて装着してください。



②パソコンと接続しましょう

SR−MOD100にフレーズを学習させたり、管理を行うためにはパソコンとの接続が必要です。パソコンでは専用のユーティリティソフトウエアで各種操作を行います。

パソコンとの接続には、当方から販売している"USB−TTL(UART)パソコン接続ボード"(型式:60R800)をご使用になると大変便利で簡単です。

60R800は、パソコンのUSBポートに接続して、仮想COMポートドライバーにより直接パソコン上からコマンド等を送信することができます。こちらの製品と、別売の7ピンヘッダーピン(型式:63CT1)を使用することで、パソコンのUSBポートに簡単に接続することができます。
またこの60R800は、+5Vの電源をUSBバスパワーから取り出すことができるため、そのままSR−MOD100に接続して電源の給電が可能です。回路を自作されない場合には、ぜひこちらのボードをご利用ください。60R800とSR−MOD100との接続方法は下記の手順で行います。

最初に60R800とパソコンを接続し、60R800のマニュアルに従い、デバイスドライバー及び仮想COMポートドライバーをインストールして、使える状態にしておきます。

60R800を使用しない場合には、ADM3202等のUART信号とRS232C信号とを変換できるレベル変換ICなどを介してパソコンと接続できるようインターフェイス回路を自作してください。参考回路図を次に掲載します。FT232R等のUART−USB変換ICなどを使うと、USBポートに接続することもできます。

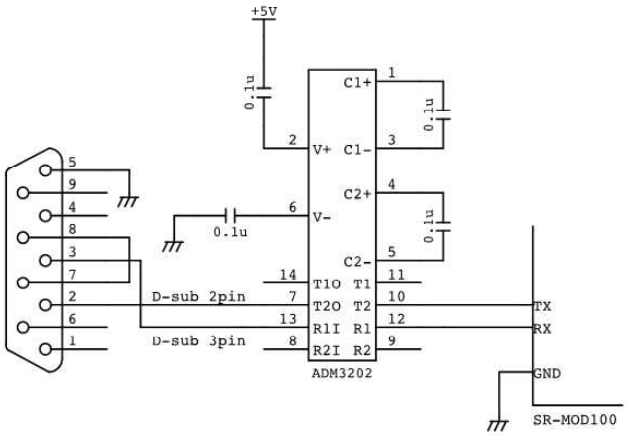
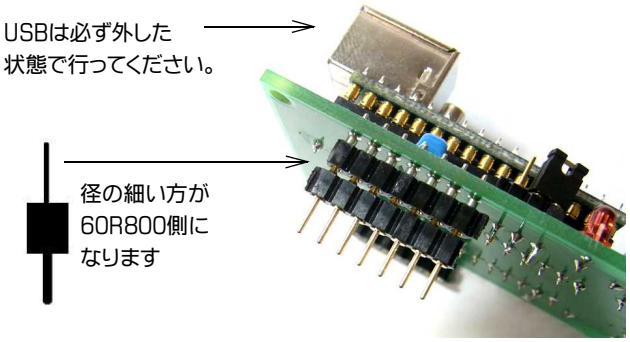
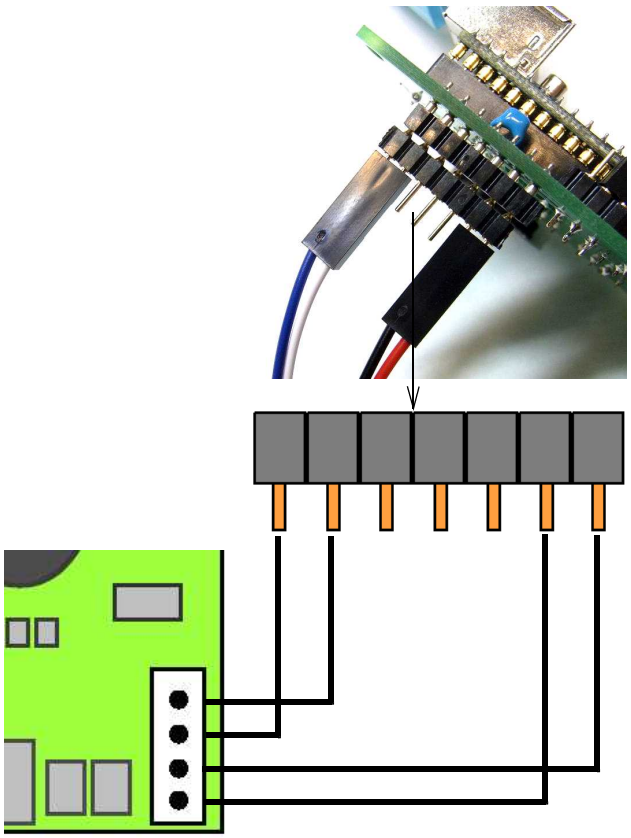


図: RS232Cインターフェイス回路参考図

1 60R800の7ピンメスヘッダーソケットに、オス−オスの7ピンヘッダーピン(型式:63CT1)を装着します。
この時、ピンの径が細い方を60R800のソケットに挿入します。



2 1で取り付けたヘッダーピンに、SR−MOD100Rに付属していたバネスを使用してSR−MOD100Rを接続します。



60R800のヘッダーピンの一番左側のピンは、SR−MOD100Rの3ピン(RX)へ、左から2番目のピンは4ピン(TX)へそれぞれ接続します。また、60R800のGNDと+5V出力ピンは、それぞれSR−MOD100Rの1ピン(GND)及び2ピン(電源)に接続します。

ハーネスは、2mmピッチ側がSR−MOD100R側に、2.54mmピッチ側が60R800のヘッダーピン側になります。

接続時は、60R800はUSBポートから必ず外しておいてください。
接続後は配線の間違いがないかをよくご確認ください。

60R800のロジック電圧設定ジャンパ(JP2)は"5.0"にセットしてください。

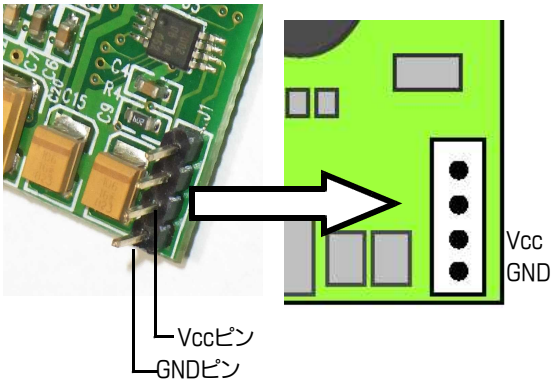
3 配線を確認後、60R800のUSBポートとパソコンのUSBポートを接続します。電源はUSBバスパワーから給電されます。
パソコンからは、仮想COMポート経由でSR−MOD100にアクセスできます。仮想COMポートの番号はデバイスマネージャで確認しておきます。

60R800をご使用にならない場合には下記の手順で電源と、UARTインターフェイスを接続します。60R800をご使用の場合には、この手順は飛ばしてお読みください。

■電源を接続しましょう

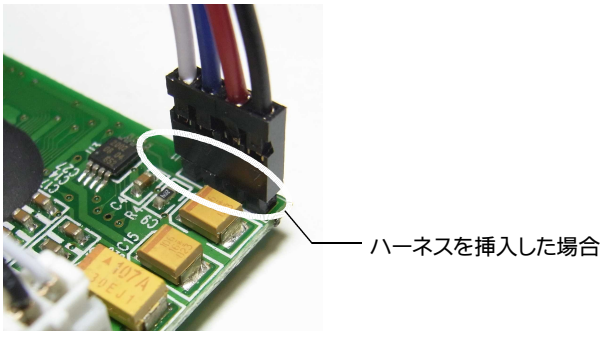
電源をSR−MOD100に接続しましょう。SR−MOD100の電源電圧は+5V又は+3.3Vになります。+5Vを印加時の消費電流は約15mA程度となります。十分に余裕のある電源をご使用ください。

電源は、SR−MOD100の電源ピンに印加します。+/−の極性を間違えるとすぐに破損しますので、接続には十分ご注意ください。



上図の通り、インターフェイスピンを右下になるように基板を置いたとき、ピンの一番下がGNDピン、その一つ上がVccピンとなります。

ピンのピッチは2mmピッチとなります。付属の赤黒ハーネスを使用すると2.54mmピッチに変換することができます。

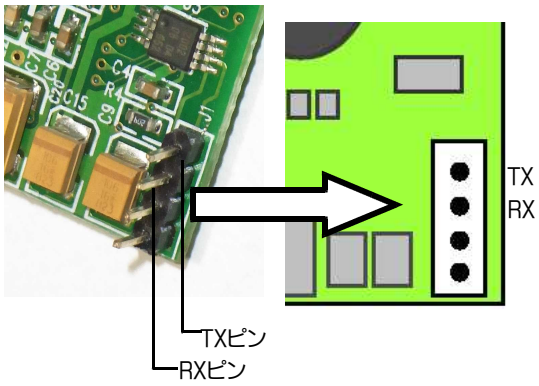


※SR-MOD100では電源投入直後は自動的にスリープモードになっています。コマンドを送信すると、ウェイクアップします。電源投入直後は自動的にスリープモードになっていることにご注意ください。

■シリアルインターフェイスを接続しましょう

SR-MOD100を駆動するためのシリアルインターフェイスを接続しましょう。通信は非同期式シリアル通信(UART)で、TX線とRX線の2線で通信を行います。CTS/RTSのフロー制御は行いません。

SR-MOD100を制御するためのホスト機器は、パソコンやUART通信のできるマイコンが接続できますが、認識させる音声のフレーズ学習や、登録等はパソコンから専用のユーティリティソフトウエアで行う必要があります。最初にパソコンのRS232Cポートと接続させるように接続されることをおすすめします。



上図の通り、インターフェイスピンを右下になるように基板を置いたとき、ピンの一番上がTXピン、その一つ下がRXピンとなります。

ピンのピッチは2mmピッチとなります。付属の青白ハーネスを使用すると2.54mmピッチに変換することができます。

パソコンと接続する場合には、当方で販売しているパソコン接続キット(型式60R800)をご使用になるのがもっとも簡単ですが、回路を自作される場合には、MAX232等のRS232Cレベル変換ICを介して、パソコンのRS232Cポートと接続してください。SR-MOD100側のUARTインターフェイスは、TTLレベル(0V-5V)です。

マイコンと接続する場合には、同じロジック電圧レベルの回路と直結できます。なお、ロジック電圧は電源電圧を絶対に超えないようにご注意ください。

ソフトウエアのインストール

SR-MOD100に音声認識フレーズを登録したり、管理を行うためのソフトウエア、"EasyVR GUI"をインストールします。本ソフトウエアは、下記サイトからダウンロードできます。

http://www.microtechnica.tv/support/software/EasyVR.zip

1 上記のファイルをダウンロードして解凍します。
"SETUP.EXE"をダブルクリックしてインストーラを起動します。

2 インストールディレクトリを指定します。デフォルトでは
"C:¥Program Files¥RoboTech srl¥EasyVRGUI"となっています。
変更する場合には直接ディレクトリを指定するか参照"ボタンを押してインストール先を設定してください。

3 "スタートメニューフォルダー"の指定はそのまま変更せず"インストール"ボタンをクリックします。インストールが開始されます。

4 インストールが完了したら"完了"ボタンを押して完了します。

SR-MOD100の認識音声の構成

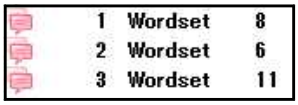
SR-MOD100には、ユーザーが自由に登録できる音声フレーズと、あらかじめ登録されているプレインストールフレーズがあります。

■プレインストールフレーズ(不特定話者認識)

プレインストールフレーズは、よく使用されるフレーズ25個で、本体にあらかじめ登録されておりユーザーは変更することができません。使用可能言語で、それぞれ25個ずつ登録されています。主にロボットの駆動用の言葉や、数字といったよく使われそうな単語が登録されています。なお数字については"4"は"シ"ではなく"ヨン"と発音する必要があります。詳しくは下記の表に記載されています。

プレインストールフレーズは、ユーザーにかかわらず、誰の声でも認識することができる不特定話者認識です。これを話者に依存しないことからSpeaker Independentとよび、"SIフレーズ"と記載します。

EasyVR GUIでは、"Group List"の中にピンク色のアイコンで表示されます。登録後は日本語の場合下記の通りです。



	Index	フレーズ
Wordset1(8フレーズ)	0	アクション
	1	進め
	2	曲がれ
	3	走れ
	4	見ろ
	5	攻撃
	6	止まれ
	7	こんにちは
Wordset2(6フレーズ)	0	左
	1	右
	2	上
	3	下
	4	前
	5	後ろ
Wordset3(11フレーズ)	0	ゼロ
	1	イチ
	2	ニ
	3	サン
	4	ヨン
	5	ゴ
	6	ロク
	7	ナナ
	8	ハチ
	9	キュウ　又は　ク
	10	ジュウ

■グループ1～15(特定話者認識)

グループ1～15までの、15のグループはユーザーが任意の認識用音声フレーズを登録できるグループとなっています。グループ分けはされていますが、ユーザーが登録できるフレーズは(話者特定パスワードフレーズ(話者識別認識)と合計して)最大で32フレーズまでです。1つのグループに32フレーズを登録することもできますし、音声の内容によって、グループ分けをすることで、誤認識などを防止できます。

SR-MOD100の音声認識開始のコマンド(0x64)は、引数としてどのグループ(1～15)に登録されたフレーズを認識させたいかを指定します。よって、指定したグループ以外のグループに属するフレーズは認識しません。グループ割けをすることで、フレーズ内容が似ており、誤認識しそうな単語であっても、場面に応じたグループ分けにより誤認識を低下させることができます。

グループ1～15に登録できる合計32フレーズは、フレーズの音声を登録学習した話者の声でのみ音声認識ができる特定話者認識です。登録学習話者以外の声では正しく認識できません。(声質が似ている場合には認識できる場合もありますが、正しい使い方ではありません。)

例えば、話者Aが「こんにちは」と登録してフレーズの学習をさせた場合、別の話者Bが「こんにちは」と発声してもSR-MOD100は認識できません。すなわち、音声フレーズを登録・学習した話者と、同じ話者が発声した時のみ音声を認識することができます。このことから話者に依存するため、Speaker Dependとよび"SDフレーズ"と記載します。

> **特定話者認識フレーズの登録数についての注意**
>
> SDフレーズは最大で32フレーズが登録できますが、これは次に紹介するパスワードフレーズ(SVフレーズ)の5つを除いた登録数です。SVフレーズを5つ登録した場合にはSDフレーズは最大27フレーズとなりますのでご注意ください。、

■話者特定パスワードフレーズ(話者識別認識)

話者特定パスワードフレーズは、最大5人までの話者を登録して、同じフレーズでも話者を識別することができる特別な認識グループです。

先のSDフレーズでは、フレーズ登録学習させた話者の声で入力された音声のみ識別可能なのに対し、本パスワードフレーズは、同じフレーズ(異なるフレーズでも可能)を最大5人まで別の話者が登録できます。登録した話者の声質をSR-MOD100は特徴として認識していますので、どの話者が発声した音声なのかを識別することができます。

例えば、「解錠(カイジョウ)」というフレーズを5名が話者としてそれぞれ学習登録させた場合、同じ「カイジョウ」というフレーズを音声認識させた時、どの話者が発声したものなのかを識別し、対応した番号をシリアル通信で返します。

この記述は、話者を特定できることからSpeaker Verificationとよび、"SV"と表記します。パスワードなどに使うことができます。

※SV機能の識別精度は、完全な本人認証や照合を保証するものではなく、信頼性はホビーレベルとお考えください。信頼性が高く求められる用途、例えば金銭や人命、財産等を保護するシステムへの使用はできません。本機のSV機能を使用されたこと、または使用できなかったことによるいかなる障害・損害についても、当方は一切の責任を負いかねますのであらかじめご了承ください。

※SR-MOD100は、SDフレーズとSVフレーズ合わせて最大32フレーズの登録ができます。SVフレーズを5フレーズ登録した場合、SDフレーズは最大27フレーズとなります。

特定話者認識(SD)グループへのフレーズ登録と学習

パソコンにインストールしたEasyVR GUIを使って、SR-MOD100に音声認識フレーズの登録をします。SR-MOD100は、あらかじめ登録された音声フレーズを内部の不揮発性メモリーに記憶します。この時、どのフレーズが「どのグループ」の「何番」かということを、フレーズ登録しながら、EasyVR GUIで管理します。

UARTでコマンド送信で行うこともできますが、登録作業に関しては、どのフレーズをどのグループの何番に登録したか・・ということがGUIで確認できるEasyVR GUIをご使用になることをお勧めします。

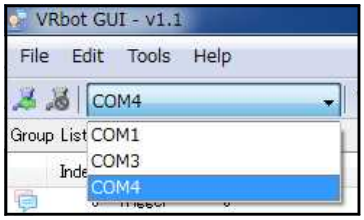
SR-MOD100では、ユーザーが任意のフレーズを32個まで登録することができます。

本項では、EasyVR GUIを使って、SR-MOD100の"グループ1～15"のグループに音声認識フレーズの登録をする方法を紹介します。なお、グループ1～15に登録できる音声認識フレーズは、SDフレーズですのでフレーズの登録学習をさせた話者と、一致する話者でないと認識はできません。

パソコンのRS232CポートとSR-MOD100を接続して使用できる状態にしてください。当方の60R800を使用されている場合には、本書の2ページ～3ページをお読みの上、正しく接続してください。

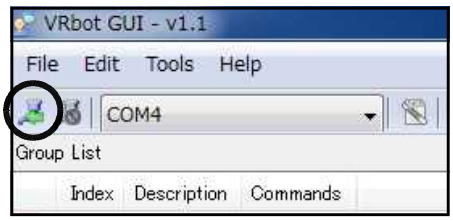
1 EasyVR GUIを起動します。

2 最初にSR-MOD100が接続されているCOMポートを指定します。プルダウンからSR-MOD100が接続されているCOMポートを選択してください。



60R800をご使用の場合には、仮想COMポートのCOM番号を指定します。COM番号はデバイスマネージャで確認できます。

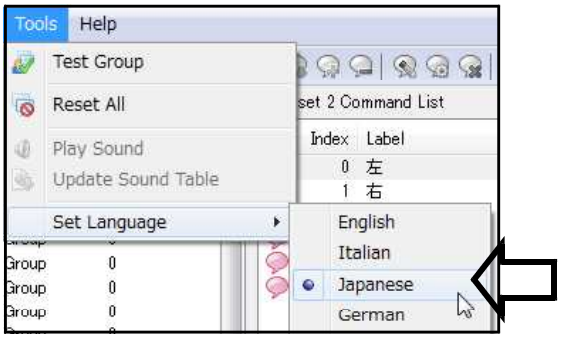
3 選択したら、"Connect"ボタンをクリックします。



正しくSR-MOD100と通信できないと下図のようなエラーが表示されます。その場合にはCOMポートの設定が正しいか、配線が正しいか等チェックをしてください。

※何度も失敗する場合には一度、SR-MOD100の電源を入れ直してください。また、60R800をご使用の場合にはJP2の設定が5.0V側に設定されているかご確認ください。

4 最初に言語を設定します。ここでは日本語を設定します。
メニューバーの"Tools"をクリックし、"Set Language"から"Japanese"をクリックして選択します。



5 EasyVR GUIでは、画面左側にフォルダが表示されており、仕分けされています。青色フォルダが特定話者認識(SD)グループで15のグループがあります。
緑色フォルダはSVのパスワード、ピンク色フォルダは、あらかじめ認識フレーズがプレインストールされている不特定話者認識(SI)グループです。

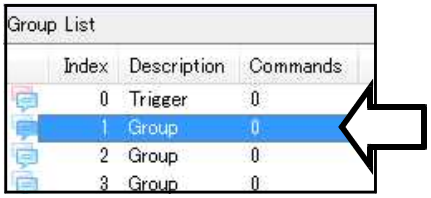
今回は例として、SDグループ(青色フォルダ)の最初のGroupに5つの下記のフレーズを登録させ認識させる設定を行いましょう。

・Index0: おはようございます
・Index1: こんにちは
・Index2: こんばんは
・Index3: コントロール
・Index4: 今週の予定

各グループは、それぞれIndexとして0から順番にインデックス番号を割り当てます。
音声認識の際には、音声がマイクから入力されて音声認識処理が行われ、該当するフレーズが登録されていると、このインデックス番号をUART経由でSR-MOD100は出力します。ホスト機器はこの値を見ることで、どのフレーズを認識したのか判断できます。

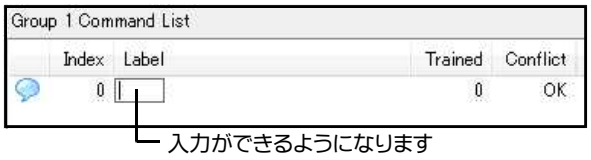
Index1～4は、フレーズの最初に"こん"が付く少しまぎらわしいフレーズとしてみました。音声認識の精度をこれで確かめます。

6 画面左側のGroup Listの一覧にある上から2番目の、"1 Group"をクリックして選択します。

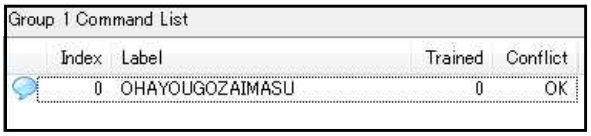


7 5つのフレーズを登録します。ツールバーの"Add command"ボタンを押します。

8 右側の"Group 1 Command List"の一覧の一番上に、Index 0として、Labelが入力できる状態となります。Labelには、フレーズを記入します。



ここでは、「おはようございます」を登録します。EasyVR GUIでは日本語テキストには対応していないため(音声認識は対応していますが表示上の仕様です)、アルファベットで "OHAYOUGOZAIMASU"と入力します。



もし、誤って入力してしまって、確定してしまった場合には、再度Labelをクリックして選択して、"Rename command"ボタンをクリックすると、再編集ができます。

9 再度、ツールバーの"Add command"ボタンを押して、手順8と同様にして、Index1～4までを追加します。

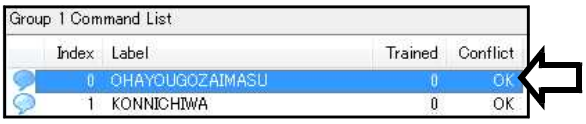
上図のように5つ登録します。
ここで"Trained"の所を見ます。現在はすべて0となっています。この回数は学習回数で、話者がこのフレーズに対して何回認識学習をさたかということを示しています。通常SR-MOD100は、それぞれのフレーズにおいて2回ずつ学習しますので、すべて学習が完了すると、ここの値はすべて2となります。

10 ではIndex0から順番にフレーズを学習させていきましょう。このグループはSDですので、特定話者認識です。認識学習をさせた話者の声でのみ認識が可能となりますので、学習する話者を決めておきます。

最初に、Index0の「おはようございます」から学習させます。SR-MOD100に付属のマイクが接続されていることを確認してください。
音声はなるべく、はっきりと通常の会話と同程度の速度で発声します。(ニュースのアナウンサーのような感じがベストです)
声の大きさは特に問いませんが、一般的な会話程度の音量とします。小さすぎる声、大きすぎる声は認識の精度が悪くなります。
また、背景雑音が多い騒がしい所や、テレビやラジオの音声が流れている所では、雑音を拾ってしまうため認識精度が著しく悪くなります。必ず静かな背景雑音のない場所で行ってください。

学習は1つのフレーズで2回行います。学習は失敗したら何度でも再学習が行えます。

Index0の"OHAYOUGOZAIMASU"の欄をクリックして選択します。

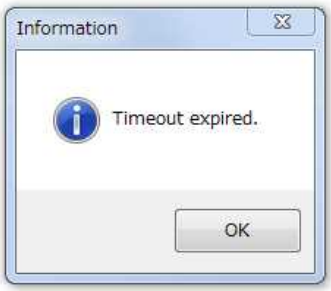


ツールバーの"Train command"ボタンをクリックします。

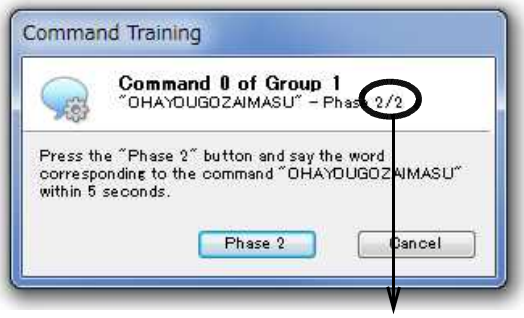
下図のようなダイアログが表示されます。

マイクに発声できるように準備ができたら、"Phase1"ボタンをクリックし、5秒以内に「おはようございます」とマイクに向かって発声します。
5秒程度無音が続くと、学習操作は自動的に完了します。

音声が正しく入力されなかった場合には、下図のようなダイアログが表示されます。この表示は音声入力がなく5秒以上経過した場合に表示されます。この場合正しく学習ができていませんので、やり直す必要があります。

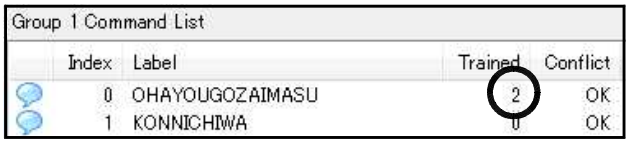


11 1回目の学習が成功すると2回目の学習を始めるための下図のダイアログが表示されます。



2回目の学習なのでフェーズが2/2となっています。
先ほどと同じように、"Phase2"ボタンをクリックしたらマイクに向かって「おはようございます」と発声します。

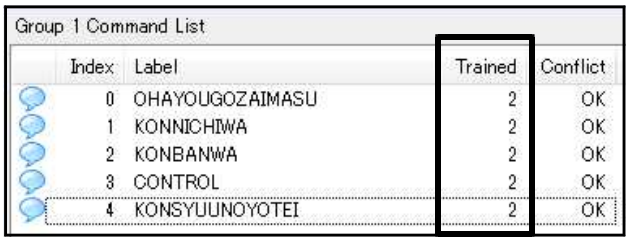
正しく学習が完了すると、ダイアログが消えて、先ほどのリスト部の"Trained"の欄に学習回数の"2"が表示されます。



12 同様にして、Index1～4の「こんにちは」～「今週の予定」までの学習を行ってください。学習はそれぞれ先ほどと同じ要領で2回ずつ行います。学習したいフレーズを選択して、"Train command"ボタンをクリックして学習を行います。

もし失敗したり、うまく学習ができていない場合には、学習内容を削除することができます。ツールバーの"Erase training"ボタンをクリックすると、学習内容は削除されます。

すべて学習が完了すると、下図のようにすべてのフレーズで"Trained"の項目が2と表示されます。



これで音声認識フレーズの登録と学習は完了しました。
続いて、登録したフレーズが正しく認識されるかを試してみましょう。

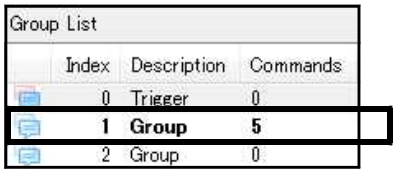
すべてのデータ(フレーズと学習内容)はSR-MOD100の不揮発性メモリーに記憶されていますので、終了したい場合には、そのままVRbot GUIを終了することができます。
また、SR-MOD100のメモリーは不揮発性ですので、そのまま電源を切断しても、データはすべて記憶されています。

特定話者認識(SD)グループの音声認識

前の項で登録・学習させた各フレーズが正しく認識させるかを実際に試してみましょう。特定話者認識(SD)ですので、認識は登録学習をさせた話者と同じ話者が発声しないと認識しません。

1 EasyVR GUIを起動し、COMポートの設定をした後、ツールバーの"Connect"ボタンをクリックして、SR-MOD100と接続します。
※詳細は5ページ手順1～3を参照してください。

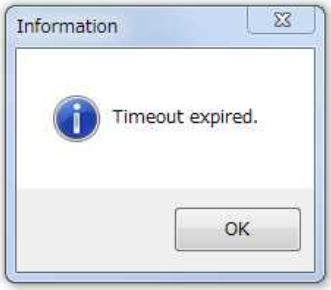
2 接続が正しく完了すると、SR-MOD100に登録されているすべてのデータが、読み込まれEasyVR GUIに反映されます。
前の項目で登録・学習させたSDグループのIndex1 Groupのコマンド数が5になっていることを確認します。コマンド数が5というのは、登録されているフレーズが5つあるという意味です。



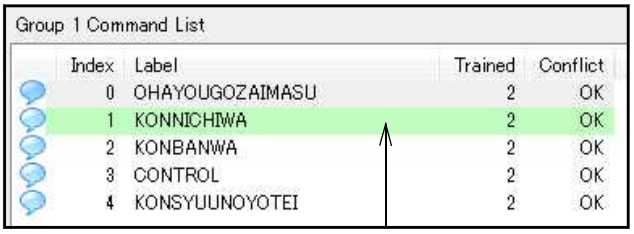
ここをクリックすると、右側の一覧に登録済みのフレーズが表示されます。

3 音声が正しく認識できるかをテストします。
マイクに話しかける準備をした後、ツールバーの"Test Group"ボタンをクリックします。

"Speak now"という表示が出ますので、ここではテストとして「こんにちは」と発声します。
ボタンを押してから5秒以内に発声してください。5秒以上無音時間が経過すると、下図のようなメッセージが表示されます。



4 正しく認識されると、Index1の"KONNICHIWA"の欄が緑色に点滅します。



緑色でハイライト点滅します

音声が入力されていても正しく認識できない場合には下図のような認識失敗の警告が表示されます。

間違ったフレーズが入力されたり、背景雑音の多い環境などで正しく登録・学習させたフレーズと一致する項目がない場合に表示されます。学習者と同一の話者であれば、認識に失敗することは希ですが、声質や"なまり"などにより認識精度が低くなることがあります。

何度も失敗する場合には再度前の項を参考にして、フレーズの学習をし直してお試しください。

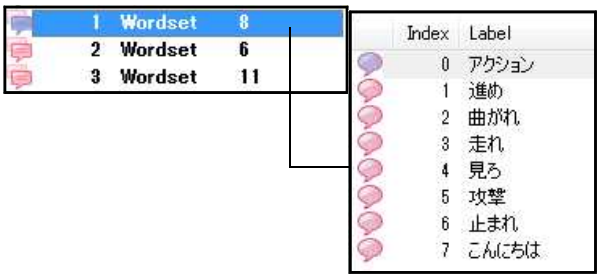
5 同様にして、5つすべてのフレーズが正しく認識できるか確認してみましょう。「こんばんは」、「今週の予定」・・・などど発声してみて正しく認識ができることを確認します。

不特定話者認識(SI)グループの音声認識

SR-MOD100には、不特定の話者の音声を認識する不特定話者認識(SI)フレーズがあらかじめ25フレーズ登録されています。このフレーズは、変更することはできませんが、不特定の話者が発声した音声を認識できますので、様々なアプリケーションに応用できます。

1 EasyVR GUIを起動し、COMポートの設定をした後、ツールバーの"Connect"ボタンをクリックして、SR-MOD100と接続します。
※詳細は5ページ手順1～3を参照してください。

2 左側の"Group List"の所にあるピンク色フォルダ3つが、不特定話者認識フレーズです。ここでは、"1 Wordset 8"をクリックします。
右側に登録されているフレーズの一覧が表示されます。



3 ここでは例として「走れ」と発音して、正しく認識するかを確認します。マイクに話しかける準備をした後、ツールバーの"Test Group"ボタンをクリックします。

"Speak now"という表示が出ますので、「はしれ」と発声します。
ボタンを押してから5秒以内に発声してください。
正しく認識されると、Index3の"走れ"の欄が緑色に点滅します。

背景雑音の多い環境などでは正しく認識されないことがあります。また、発声に"なまり"があったりすると認識精度が落ちます。

その他のフレーズについても、お試しください。
"3 Wordset 11"にはゼロ～ジュウまでの11個のフレーズが登録されていますが、下記の数字は発音に注意してください。
・4 → ヨン （「シ」は認識しません）
・7 → ナナ （「シチ」は認識しません）

不特定話者認識フレーズの誤認識が多い場合

不特定話者認識フレーズの誤認識が多い場合には、言語設定が誤っている場合があります。その場合には、一度ツールバーの"Set Language"ボタンをクリックして、"English"を選択し、言語を一度英語に設定した後、再度、"Set Language"ボタンをクリックして、"Japanese"を選択してお試しください。

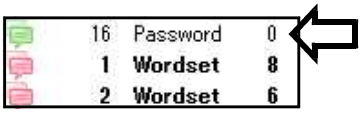
話者特定認識(SV)グループの音声認識

話者特定パスワードフレーズ(SV)は、最大5人までの話者を登録して、同じフレーズでも話者を識別することができます。同じフレーズで話者を識別できるので、発声話者に応じた別の処理を行うアプリケーションなどに応用できます。

ここでは、例として3人の話者が「こんにちは」という同じフレーズを登録し、それぞれ識別されることを確認します。

1 EasyVR GUIを起動し、COMポートの設定をした後、ツールバーの"Connect"ボタンをクリックして、SR-MOD100と接続します。

2 左側の"Group List"の所にある緑色フォルダが、話者特定認識フレーズです。"16 Password"をクリックします。



3 フレーズを登録します。「こんにちは」というフレーズを話者毎に用意します。ツールバーの"Add command"ボタンを押します。

4 右側のLabelの部分が入力待機状態となりますので、"KONNICHIWA_1"と入力します。同様にして、下図のように3つ登録をしてください。話者が特定できるように、"_1"など識別できる番号を付けておいてください。

Index	Label	Trained	Conflict
0	KONNICHIWA_1	0	OK
1	KONNICHIWA_2	0	OK
2	KONNICHIWA_3	0	OK

上図のように3つ登録します。

5 最初に話者1が、「こんにちは」を登録学習させます。
一番上の "0 KONNICHIWA_1" の部分をクリックして選択しハイライト表示させます。

Index	Label	Trained	Conflict
0	KONNICHIWA_1	0	OK
1	KONNICHIWA_2	0	OK
2	KONNICHIWA_3	0	OK

この状態で、"Train Command"ボタンをクリックします。

6 下図のようなダイアログが表示されますので、マイクに向かって発声する準備ができたら、"Phase1"ボタンをクリックします。"Speak Now"と表示されますので、5秒以内に「こんにちは」と発声します。

7 同様にして、2回目の学習も行いますので準備が整ったら"Phase2"ボタンをクリックして、「こんにちは」と発声します。
正しく学習が完了すると、リストの"Trained"の数値が2になります。

Index	Label	Trained	Conflict
0	KONNICHIWA_1	2	OK
1	KONNICHIWA_2	0	OK
2	KONNICHIWA_3	0	OK

2回登録学習が完了したことを表します。
同様にして、話者2及び話者3も、音声フレーズの学習をさせてください。3人の話者が学習完了すると、"Trained"の数値はすべて2となります。

Index	Label	Trained	Conflict
0	KONNICHIWA_1	2	OK
1	KONNICHIWA_2	2	OK
2	KONNICHIWA_3	2	OK

8 続いて、正しく話者を識別できるか確認します。
ここでは例として話者1の人が、「こんにちは」と発音して、正しく1番目の"KONNICHIWA_1"が認識されるかを確認します。
マイクに話しかける準備をした後、ツールバーの"Test Group"ボタンをクリックします。

"Speak now"という表示が出ますので、話者1が「こんにちは」と発声します。正しく "KONNICHIWA_1" が緑色にハイライト点滅表示されれば成功です。

Index	Label	Trained	Conflict
0	KONNICHIWA_1	2	OK
1	KONNICHIWA_2	2	OK
2	KONNICHIWA_3	2	OK

同様にして話者2、話者3でも正しく識別されるかどうか確認ください。
ほとんどの場合、SR-MOD100は話者の声質を特徴として分析学習しているため、識別できますが、希に声質が似ていたりする場合には、正しく識別ができない場合もあります。SR-MOD100の識別精度は高いですが、識別精度を保証するものではありませんので、あらかじめご了承ください。

シリアルコマンドでの制御について

SR-MOD100の駆動に使用されるシリアルコマンドは、非同期式シリアル通信(UART)で、フロー制御を使用しないTXとRXを接続する2線式の通信をホスト機器と行います。

ロジック電圧レベルは、電源電圧を最大値として、3.3V～5.0VのTTLレベルとなっています。

電源投入時の初期状態では、通信速度9600bps、データ長8ビット長、1ストップビット、ノンパリティの設定になっています。通信速度は、シリアルコマンドで最高115.2kbpsまで変更させることができます。
本製品の場合、速度は必要としないため通常は9600bpsで使用することをお奨めします。PICマイコン等のTTLレベルのUARTインターフェイスと接続する場合には直結できます。パソコンのRS232Cポートと接続する場合には、別途MAX232等のレベル変換ICを介して接続します。

■数値表現について

SR-MOD100の数値表現は少し特殊な方法を採用しています。引数や戻り値で使用される整数値は、-1～31まで用意されていますが、それらは、「40hを最小値の-1」とし「60hを最大値の+31」としている点です。よって、整数値0は、41hとなります。

整数値　最小値	
40h	整数の最小値　-1

整数値　最大値	
60h	整数の最大値　+31

整数値　0	
41h	整数のゼロ　0

よって、例えば言語を指定するコマンド(CMD_LANGUAGE)は、コマンドが6Chで、英語に設定する場合引数は0です。この場合、0という引数は0を送信するのではなく、41hを送信することになります。よって、言語を英語にする場合には、6Ch,41hををシリアル経由で送信することになります。最も簡単な算出方法は、41hを送信したい値に加算することです。例えば、"5"という数値を送信したい場合、SR-MOD100には46hを送信します。よって5+41="46"→46hを送信すればよいことにります。同様にして、最大値"31"ならば31+41h=60hとなります。

逆に受信した値から数値をデコードする場合には、41hを減算すればよいことになります。受信した値が5Bhだとしたら、5Bh-41h=26なので26ということになります。

Hex	40h	41h	42h	43h	44h	45h	･･･	5Eh	5Fh	60h
値	-1	0	1	2	3	4	･･･	29	30	31

■SR-MOD100からのAcknowledgement(ACK)について

ACKnowledgement(ACKと記載)は、SR-MOD100がコマンドを受信した時、そのコマンドが正しく受信されたことを通知するために出力する肯定的な返事のデータです。SR-MOD100のACKは6Fhです。

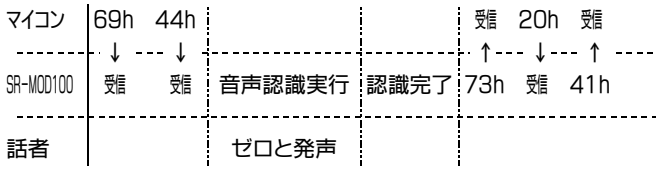
あるコマンドを送信してその内容にエラーが無く正しく処理が行われると、6Fhが戻りますので、次の処理は必ずこのACKを受信してから行うようにします。なお、このSR-MOD100からの戻り値は初期設定では、20ミリ秒後に戻るよう設定されています。

■複数の戻り値があるコマンドの場合

戻り値が2つ以上あるコマンドの場合には、SR-MOD100は1つずつ戻り値を戻し、次の戻り値をSR-MOD100から送信させたい場合には、必ず"ARG_ACK"(20h)コマンドを送る必要があります。"ARG_ACK"をホスト側から送信しないと、次の戻り値をSR-MOD100は送信できませんので、ご注意ください。

例えば、不特定話者音声認識が正しく完了した場合、SR-MOD100は完了したことを通知する"STS_SIMILAR"(73h)と、それに続いて該当するインデックス値を返します。この時、SR-MOD100は最初の73hを送信後、"ARG_ACK"(20h)を受信するまでインデックス値の送信を待機します。"ARG_ACK"(20h)を受信後、インデックス値を返します。

例:不特定話者認識を実行したワードセット3に対して実行した場合



■ウェイクアップ信号について

SR-MOD100では電源投入直後は自動的にスリープモードになっています。よって電源投入初回にコマンドを送信した場合、"STS-AWAKEN"(77h)がSR-MOD100から戻り、ウェイクアップします。この77hが戻った場合には、その直前にSR-MOD100が受信したコマンドやデータは破棄されますので、あるコマンドをSR-MOD100に送信後、77hが戻った場合には、そのコマンドをもう一度送信する必要があります。

SR-MOD100にコマンドを送信する場合には必ず、コマンド送信後に77hが返らないかを確認し、77hが返った場合には、再度そのコマンドを送信するようホスト側のプログラムを設計して頂けますようお願い致します。特に電源投入直後は常にスリープモードになっていることに留意してください。

※電源投入直後以外でも、スリープモードになっている場合がありますので、上記の通りホスト側からコマンドをSR-MOD100側に送信する場合には必ず、77hが返らないかをご確認ください。

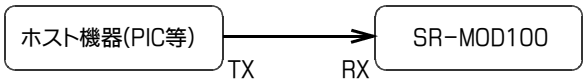
■使用言語設定についてのご注意

SR-MOD100では使用言語を設定するコマンド、"CMD_LANGUAGE"(6Ch)コマンドがあります。本コマンドによって、SR-MOD100が使用する言語を設定しますが、その設定は電源を切断すると、リセットされ不定の状態となります。(不揮発性メモリーに記憶されません)
よって、電源投入後には必ず使用言語の設定を行う必要があります。本設定を行わないで使用すると、認識精度が著しく悪くなります。よって、認識精度を確保するために、電源を投入後は本コマンドを必ず実行してから各種コマンドを送信するようにホスト側の設計をして頂けますようお願い致します。

なお、電源投入後に言語設定コマンド、"CMD_LANGUAGE"を実行するという処理が難しい場合には、、"CMD_RECOG_SD"・"CMD_RECOG_SI"・"CMD_TRAIN_SD"のコマンドを実行する前に、必ず本コマンドを実行するように設計されると安全です。見落としがちな設定項目ですので、十分ご注意ください。

■シリアルコマンドの時間間隔について(ホスト→SR-MOD100)

シリアルコマンドは、1バイトだけの場合(コマンド本体のみ)と複数バイトの場合(引数を伴う場合)があります。引数を伴う複数バイトのデータをSR-MOD100に送信する場合には、1バイト毎に約20ミリ秒程度の遅延を挿入して頂けますようお願い致します。

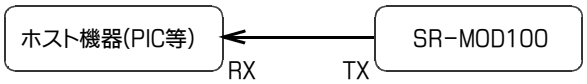


例えば言語設定コマンドCMD_LANGUAGE(6Ch)の場合、引数に設定言語を指定します。日本語の場合には続いて43hを引数として送信します。この時、6Chと43hの間隔を20ミリ秒程度開けて下さい。

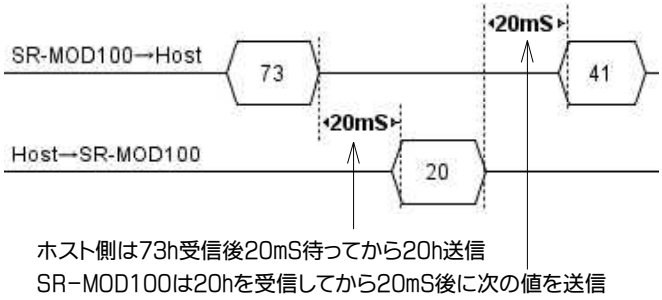
■シリアルコマンドの時間間隔について(SR-MOD100→ホスト)

続いて、SR-MOD100から返される戻り値についてですが、戻り値についても複数バイトの戻り値がある場合には各バイト毎にディレイが設定されています。デフォルト設定では20ミリ秒となっています。このディレイ値はコマンドによって0～10ミリ秒、20～100ミリ秒、200～1000ミリ秒の間でそれぞれ設定できます。通常はデフォルトの20ミリ秒を使います。

SR-MOD100では複数の戻り値を返す場合1バイト毎に次の戻り値を要求する20hを受信するホスト側が送信する必要があります。この20ミリ秒の遅延というのは、SR-MOD100が20hをホストから受信した時に次の戻り値を返すまでの時間です。



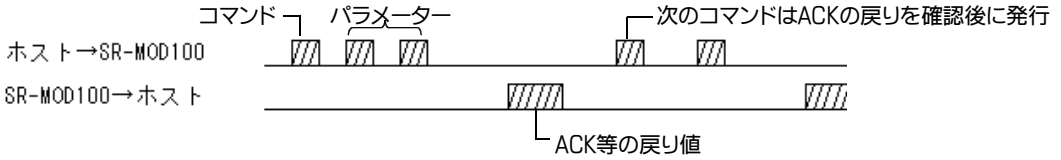
例えば、不特定話者認識を実行した結果、認識が完了しその結果をSR-MOD100がホスト側にコマンドで通知する場合、SR-MOD100は認識完了を示す73hと、それに続いてどのワードを認識したのかということを示すインデックス値40h～60hまでを返します。ホスト側は最初の73hを受信したら、次の引数を受信するために73h受信後から約20ミリ秒程度時間を置いてから20hをSR-MOD100に送信します。
SR-MOD100はこの20hを受信してから20ミリ秒後に認識結果であるインデックス値をホスト側に送信します。



シリアルコマンド一覧

コマンドに書かれた値は、16進数です。例えば、"6Ch"と書かれたコマンドの場合には、16進数の6Cという意味になります。(C言語などでは0x6Cと記載することがあります。)コマンドはすべて1バイトで、必要に応じてパラメーター(引数)があります。

各コマンドを送信すると、コマンドの内容に応じて戻り値(SR-MOD100からホストへ戻る値)がある場合があります。戻り値がある場合には、各コマンド解説の項に「期待戻り値」の種類が記載されています。戻り値があるコマンドの場合には、必ず戻り値が戻ったことを確認した後、ホスト機器は次のコマンドを送信するように設計する必要があります。期待戻り値とは、そのコマンドが実行された時、または実行が完了した後に返ることが予想される戻り値のことをいいます。多くの場合、ACK(正常完了の通知)である"STS_SUCCESS(6Fh)"の戻りが期待されます。



ACK等、戻り値が戻る前にコマンドを送信してしまうと、正しくSR-MOD100は動作を行いませんのでご注意ください。また、場合によってはエラーコードが返っている場合がありますので、必ず戻り値を読んで、エラーの場合にはその対応をできるようホスト側プログラムを作成してください。

引数に記載されている値は、SR-MOD100の数値表現にして送信する必要があります。例えば、引数に "0"(ゼロ)と記載されている場合には、41hを送信します。"1"(イチ)と記載の場合には、42hです。ご注意ください。なお引数の項目がないコマンドには引数の設定はありません。

下記コマンド一覧の引数の部分に記載の[ ]内の値は、引数のバイト数を表します。

■SR-MOD100Rを制御するコマンド一覧

【機能/動作】 *CMD_BREAK* 音声認識の動作を中止します、又は現在の動作状態を確認します

【コマンド】 62h ('b')

【解説】 音声認識プロセスが実行中の場合には、このコマンドを送信することでSR-MOD100は認識プロセスを中止します。中止した場合には、戻り値としてSTS_INTERR(69h)が返ります。同様にフレーズ認識学習を強制的に終了したい場合も使用できます。
その他、何も動作をしていない場合で、コマンドを受け入れ可能な場合には、STS_SUCCESS(6Fh)を返します。何かコマンドを送信する前に、このコマンドを送信し、STS_SUCCESSが返ることを確認することで、SR-MOD100がコマンド受け入れできる状態かどうかを確認することができます。
その他、パワーダウンモード実行中でコマンド受信でウェイクアップをしたい場合には、このコマンドを使用することで、ウェイクアップだけを実行できます。

【期待戻り値】 STS_SUCCESS(6Fh) STS_INTERR(69h)

【機能/動作】 *CMD_SLEEP* 指定したパワーダウンモードを実行します

【コマンド】 73h ('s')

【引数】 1つ [1]スリープモードを0~8の中から1つ指定します。
0 = キャラクタを受信した時だけウェイクアップします
3~5 = 2回のクラップ音(手を叩く音)の入力(解説参照)及びキャラクタを受信した時にウェイクアップします
6~8 = 3回のクラップ音(手を叩く音)の入力(解説参照)及びキャラクタを受信した時にウェイクアップします

【解説】 SR-MOD100の消費電流は+5V電源使用時で、動作時平均約11mA程度です。スリープモードになると消費電流はかなり少なくなります。ウェイクアップの方法によって、スリープ時の消費電流が異なり、モード0の場合には待機電流は20μA程度に、その他のモードの場合には0.25mA程度になります。
引数3~5は、手を2回叩く音(クラップ)によってウェイクアップします。引数の違いは音の感度の違いで、3が最も感度が良く比較的小さな音で反応します。5を指定すると、少し大きめのクラップ音に反応します。
なお、クラップ音は2回連続で手を「パチパチ」と叩く音で、間隔は1秒~2秒程度にします。(間隔が空きすぎるとクラップ音として認識されないことがあります。)
引数6~8は、手を3回叩く音によってウェイクアップします。引数の違いは上記と同じで感度の違いとなります。3回「パチパチパチ」と手を叩くとウェイクアップします。なお両方ともコマンド受信でもウェイクアップします。

【期待戻り値】 STS_SUCCESS(6Fh)

【機能/動作】 *CMD_LANGUAGE* 使用言語を指定します

【コマンド】 6Ch ('l')

【引数】 1つ [1] 設定言語: 0=英語、 1=イタリア語、 2=日本語、 3=ドイツ語、 4=スペイン語、 5=フランス語

【解説】 SR-MOD100の使用する言語を設定します。本設定は電源投入後、最初の音声認識(SDフレーズ、SIフレーズ、SVフレーズいずれの場合も)及び最初のフレーズ学習をする前に必ず設定を1回する必要があります。
電源投入中は、その設定は保持されますが、電源切断後、再度電源を投入した場合には、初回の音声認識時並びにフレーズ学習時には、その実行前に言語設定をする必要があります。本設定を行わないで、電源投入後に音声認識をすると認識精度が悪くなる場合があります。必ず電源投入後初回の音声認識、音声フレーズ登録の時にはその作業前に本言語設定を行ってください。

【期待戻り値】 STS_SUCCESS(6Fh)

【補足】 言語設定は基本的には内部のメモリーに記憶されますが、電源切断後にはその設定がリセットされていることがあります。本設定は認識精度に著しく影響を与えますので認識精度を確保するための対策として、電源投入後には毎回設定を行って頂けますようお願い致します。なお電源投入直後は、スリープモードになってるため、電源投入直後に本コマンドを送信した場合には"STS_AWAKE"が返る場合があります。その場合には、再度送信をしてください。
もし、電源投入初回に本コマンドを確実に実行できない場合には、"CMD_RECOG_SD"・"CMD_RECOG_SI"・"CMD_TRAIN_SD"のコマンドを実行する前に、必ず本コマンドを実行するように設計されると安全です。

※本コマンドは何度実行されても問題ありません。

【機能/動作】 *CMD_TIMEOUT* 音声認識のタイムアウト時間を設定します

【コマンド】 6Fh ('o')

【引数】 1つ [1] タイムアウト時間: −1=デフォルト(3秒)、 0=無限待機、 1〜31=秒数設定

【解説】 音声認識開始コマンドを送信後、何秒間音声入力を待機するかを設定します。指定した秒数が経過すると、タイムアップして認識プロセスは終了します。デフォルト設定では3秒になっており、1〜31までを秒単位で指定できます。
なお、無限待機に設定すると、音声認識開始後からタイムアウト制限無く音声入力を待機しますが、周囲の雑音等によって音声が入力されたと認識して、正しい認識ができないで認識プロセスが終了することがありますので推奨できません。
デフォルト時間(3秒)での使用が最も推奨されます。

【期待戻り値】 STS_SUCCESS(6Fh)

【機能/動作】 *CMD_BAUDRATE* UARTの通信速度を設定します

【コマンド】 61h ('a')

【引数】 1つ [1] 通信速度: 1 = 115200 , 2 = 57600, 3 = 38400 , 6 = 19200, 12 = 9600(デフォルト設定) 単位:bps

【解説】 非同期式シリアル(UART)の通信速度を5種類から選択して設定します。
引数は1から順番ではなく3の次が6、6の次が12であることに注意してください。
設定直後に返る戻り値は、設定前の速度で戻ります。
通信速度の設定は揮発性です。電源を切断すると、設定はリセットされ9600bpsになります。

【期待戻り値】 STS_SUCCESS(6Fh)

【機能/動作】 *CMD_RESETALL* すべての設定及び登録したSD及びSVグループのフレーズを消去します

【コマンド】 72h ('r') 送信後に確認として 52h('R') を送信

【解説】 SR-MOD100に設定されているすべての設定及び、ユーザーが登録した特定話者認識フレーズ(SDフレーズ)、話者特定パスワードフレーズ(SVフレーズ)の登録を削除しリセットします。削除されたSD及びSVフレーズは復旧できません。

本コマンドは、すべての不揮発性メモリーの内容を消去するため、誤って実行しないように2つのコマンドを送信することで実行されるようになっています。最初に72hを送信後、52hを送信することで消去が実行されます。72hを送信後に、"CMD_BREAK"(62h)を送信すると本コマンドはキャンセルされます。

リセット及び消去は52hを送信後から開始され、終了までに約15秒～40秒程度かかります。リセット及び消去が完了すると、"STS_SUCCESS(6Fh)"が返りますので、6Fhが返るまで他のコマンド等を送らないように待機する必要があります。次の作業は必ず6Fhが返ったことを確認してから行うように設計しなければなりません。

【補足】 本コマンドは、内蔵メモリーを初期化することですべての機能をリセットします。予期しない電源断等によって内蔵メモリーの構造が壊れた場合も本コマンドを実行することで修復することが可能です。但しその場合には登録されている学習フレーズはすべて削除されますので、その後再度フレーズの登録や学習が必要となります。

【期待戻り値】 STS_SUCCESS(6Fh)

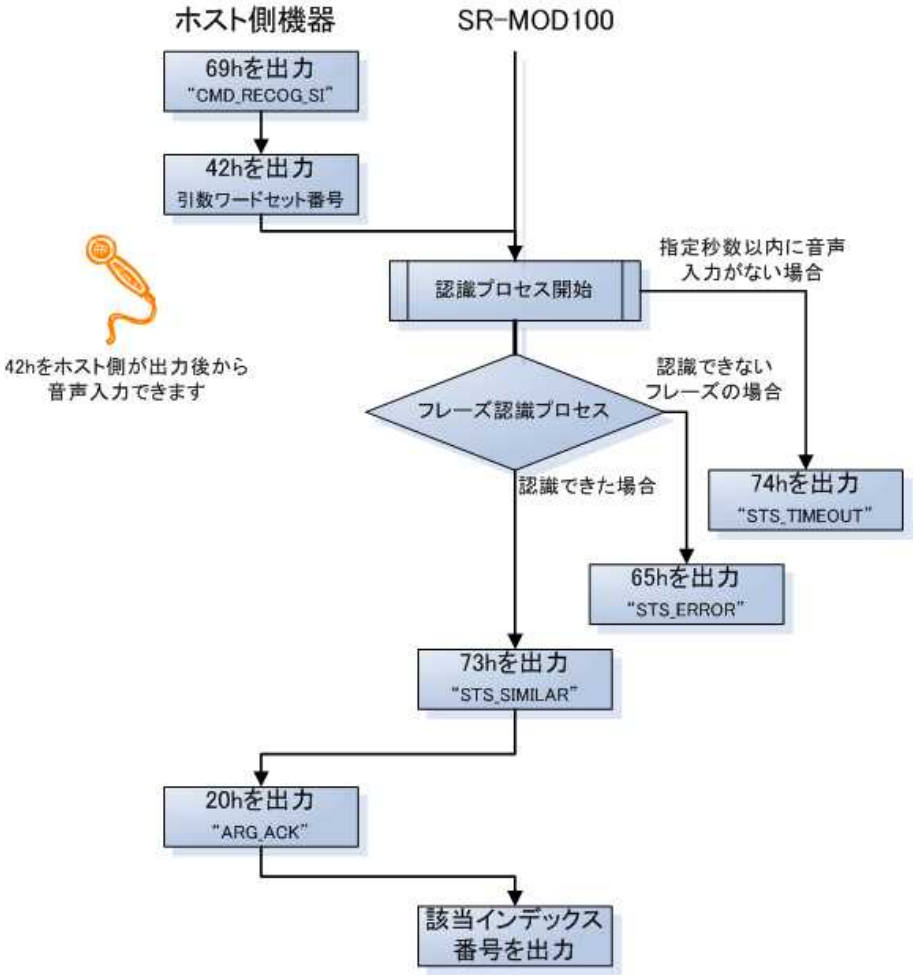
【機能/動作】 *CMD_RECOG_SI* 不特定話者認識プロセスを実行します

【コマンド】 69h ('i')

【引数】 1つ [1] ワードセットインデックス: 1～3

【解説】 不特定話者音声認識(SI)プロセスを実行します。プレインストールフレーズのワードセットはあらかじめ1～3に分類されています。どの分類の音声認識を実行するのかを引数で指定します。プレインストールフレーズ一覧は本書4ページをご覧ください。

音声認識プロセスは、引数の値を送信した時点から開始されます。タイムアップ時間を指定している場合には、引数の値送信後から音声入力がされる期間を時間カウントします。一連の動作フローを以下に記載します。
認識が正しく行われると、"STS_SIMILAR"(73h)が返り、ホスト側から"ARG_ACK"(20h)を送信すると、認識したフレーズに該当するインデックス値を返します。



上記例ではプレインストールフレーズとしてワードセット番号に1番(42h)を指定していますので、例えばこの例では「こんにちは」と発声した場合、リストのIndex番号7に該当するフレーズなので、上記フロー図の最後は48hが返ることになります。

認識に成功しリストにあるフレーズだった場合には"STS_SIMILAR"(73h)を、認識できないフレーズだった場合には"STS_ERROR"(65h)を、設定したタイムアウト時間以内に音声入力がなかった場合には、"STS_TIMEOUT"(74h)を出力します。

なお、注意が必要なのはフレーズの認識が出来た後、SR-MOD100は"STS_SIMILAR"(73h)を出力して認識できたことを通知してきますが、ホスト側が"ARG_ACK"(20h)を出力しないと、認識結果のインデックス番号を出力しないことにご注意ください。

【期待戻り値】 STS_SIMILAR(73h) , STS_TIMEOUT(74h) , STS_ERROR(65h)

【補足】 本コマンドを電源投入後に最初に実行する場合には、本コマンド実行の前に必ず"CMD_LANGUAGE"(6Ch)にて使用言語の設定を行っておく必要があります。使用言語の設定は、本来電源投入後電源が切断されなければ、1回だけでよいですが、本コマンド実行前に必ず言語設定を実行してから、本コマンドを実行するよう設計されても問題ありません。

【機能/動作】 *CMD_GROUP_SD* 特定話者認識(SD)及び話者特定(SV)フレーズをグループに新規登録します

【コマンド】 67h ('g')

【引数】 [1] グループインデックス: 1～15 = SDフレーズ 16 = SVパスワードフレーズ
[2] コマンドポジション: 0～31 = コマンドポジションの指定

【解説】 特定話者認識(SD)フレーズを登録するグループ(1～15)又は、話者を特定(SV)するフレーズを登録するグループ(16)に、新規にフレーズを登録します。
新しくSDフレーズ又はSVフレーズを作る場合には、このコマンドで任意のグループにフレーズを登録します。フレーズは、各グループにそれぞれフレーズを登録することが可能で、各フレーズの番号をコマンドポジションと呼びます。

なお、この時点では新規にフレーズを登録するエリアは登録されましたが、まだ実際にはフレーズや、フレーズの内容がわかる名前(ラベル)は登録していませんので、このあとの別のコマンドでフレーズの学習登録、フレーズに対する名前付けを行います。
なお、すでに登録可能なメモリー容量を超えて登録しようとすると、"STS_OUT_OF_MEM"(6Dh)が返り、新規登録ができないことを通知します。登録は最大で、パスワードフレーズ(SVフレーズ)と合わせて32フレーズとなります。

登録されたエリアは内蔵の不揮発性メモリーに記憶されますので、電源を切断してもその内容は保持されます。

【期待戻り値】 STS_SUCCESS(6Fh) , STS_OUT_OF_MEM(6Dh)

【使用例】 例:話者特定認識(SD)フレーズの、グループ1にコマンドポジション番号0番のフレーズを新規に登録する場合には、次のようにコマンドを送信します。

67h - 42h - 41h

【機能/動作】 *CMD_NAME_SD* 特定話者認識(SD)及び話者特定(SV)フレーズに名前(ラベル)を付けます

【コマンド】 6Eh ('n')

【引数】 [1] グループインデックス: 1～15 = SDフレーズ 16 = SVパスワードフレーズ
[2] コマンドポジション: 0～31 = コマンドポジションの指定
[3] ラベルの長さ: 1～31 = 名前(ラベル)の文字列の長さを指定します
[4]～[n]ラベルの文字列(ASCII): 名前(ラベル)をASCIIキャラクタで引数[3]で指定した文字数分送信します

【解説】 SR-MOD100では、"CMD_GROUP_SD"(67h)で新規に作成したSDフレーズ又はSVフレーズの登録エリアに対して、後からそれが何というフレーズだったかを文字列として分かるように、登録学習させるフレーズに対して名前(ラベル)を31文字以内で付けることができます。文字列はASCIIキャラクタで指定し、記号は使用できません。
例えば、わかりやすい例としては、「こんにちは」というフレーズならば、"KONNITIWA"などの文字列を登録しておきます。

登録された名前は内蔵の不揮発性メモリーに記憶されますので、電源を切断してもその内容は保持されます。

特定話者認識(SD)フレーズは、グループが1～15グループまでありますので、任意のグループを選びます。
話者を特定する(SV)パスワードフレーズは、16グループにのみ登録できます。
各グループには0～31までの計32個のフレーズを登録することができます。コマンドポジションとは、そのグループ内での番号のことをいいます。"CMD_GROUP_SD"で登録したグループインデックス内のフレーズ(コマンドポジション)に対して、名前を付けます。

【補足】 このコマンドを実行する前に、"CMD_GROUP_SD"(67h)を実行して、SDフレーズグループ又はSVフレーズグループにコマンドを作成してから実行してください。コマンドが先にできていないと、名前を登録することができません。

【期待戻り値】 STS_SUCCESS(6Fh)

【使用例】 例:SDグループのグループインデックス1のコマンドポジション0に登録されているSDフレーズに対して、"HELLO"という名前(ラベル)を付ける場合

6Eh	42h	41h	46h	48h	45h	4Ch	4Ch	4Fh
	1	0	長さ5	"H"	"E"	"L"	"L"	"O"

【機能/動作】 *CMD_TRAIN_SD* SDフレーズ・SVフレーズの認識フレーズの学習を実行します

【コマンド】 74h ('t')

【引数】 [1] グループインデックス: 1～15 = SDフレーズ 16 = SVパスワードフレーズ
[2] コマンドポジション: 0～31 = コマンドポジションの指定

【解説】 先に"CMD_GROUP_SD"(67h)で作成したSDフレーズ又はSVフレーズに対して、音声認識に必要なフレーズ学習を実行します。
話者がSR-MOD100に接続されたマイクに向かって、発声することで、その発声内容をSR-MOD100は学習して内蔵された不揮発性メモリーに登録します。

学習は1つのフレーズに対して最高6回まで行うことができます。但し、実用的には2回～3回程度の学習で十分な認識精度が得られます。学習の回数は3回より多くしても認識精度の向上は期待できませんので、多くても3回程度までをお奨めします。
なお、学習は7回以上行わないようにしてください。
声質や周囲の雑音等によって認識精度は大きく変わりますので、実際に認識の精度を確かめながら2回～3回程度の反復学習をさせることをお奨めします。反復学習させる場合には、再度本コマンドを実行することで、学習を積み重ねることができます。

学習した回数(TRAIN COUNT)は、"CMD_DUMP_SD"(70h)コマンドで確認することができます。

学習は、3バイト目のコマンドポジションのデータを受信した直後から開始されます。コマンドポジションデータをホスト側から送信したら、すぐに(3秒以内)音声をマイクに向かって発声します。終了のコマンドはなく、音声が入力されてから、無音時間が3秒経過した時点で学習は自動的に終了します。

背景雑音の多いところや、人間の音域を超える範囲の音声などでは正しく学習ができず認識精度が著しく悪くなることがあります。また、イントネーションが不明瞭だったり、滑舌のよくない発声だと認識精度に悪影響を与えることがあります。発声する際には周囲が静かな所で、はっきりとした声でマイクに向かって発声するようにします。

【補足】 同じフレーズを、同一グループインデックス内に重複して学習させようとすると、"STS_RESULT"(72h)でエラーが発生します。このエラーではフレーズが重複していると思われるコマンドポジションの値を通知します。なお、グループインデックスが異なる場合には、同じフレーズを登録することができます。

学習プロセスは音声入力後に3秒以上経過すると完了します。完了後には学習が正しく完了したか、又はエラーが発生したか必ず戻り値が戻りますので、この値を必ず確認するように設計します。エラー"STS_ERROR"(65h)が戻っている場合には、更にその詳細が戻りますので、原因を特定する目安になります。(詳しくは"STS_ERROR"(65h)の項をご覧ください。)

学習をさせるフレーズについては、予め"CMD_GROUP_SD"(67h)コマンドで、新規にフレーズを登録しておく必要があります。また、どんなフレーズかを分かりやすくするため、"CMD_NAME_SD"(6Eh)でラベルを付けておきます。

本コマンドを電源投入後に最初に実行する場合には、本コマンド実行の前に必ず"CMD_LANGUAGE"(6Ch)にて使用言語の設定を行っておく必要があります。使用言語の設定は、本来電源投入後電源が切断されなければ、1回だけでよいですが、本コマンド実行前に必ず言語設定を実行してから、本コマンドを実行するよう設計されても問題ありません。

【期待戻り値】 STS_SUCCESS(6Fh) , STS_RESULT(72h) , STS_SIMILAR(73h) , STS_TIMEOUT(74h) , STS_ERROR(65h)

【機能/動作】 *CMD_ERASE_SD* SDフレーズ・SVフレーズのフレーズ学習を消去します

【コマンド】 65h ('e')

【引数】 [1] グループインデックス: 1～15 = SDフレーズ 16 = SVパスワードフレーズ
[2] コマンドポジション: 0～31 = コマンドポジションの指定

【解説】 指定したグループインデックスの、コマンドポジションについて、"CMD_TRAIN_SD"(74h)で行った認識学習の学習内容を消去します。このコマンドを実行すると、学習は消去され学習回数は0になります。
なお、本コマンドはフレーズの登録を削除するものではなく、学習内容を消去するものですのでご注意ください。フレーズ登録そのものを削除する場合には、"CMD_UNGROUP_SD"(75h)コマンドをご利用ください。

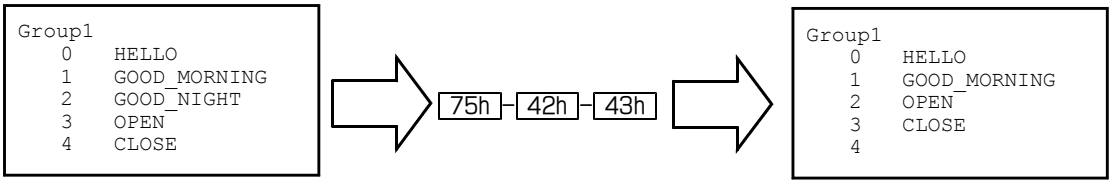
【期待戻り値】 STS_SUCCESS(6Fh)

【機能/動作】 *CMD_UNGROUP_SD* SDフレーズ・SVフレーズの登録フレーズを削除します

【コマンド】 75h ('e')

【引数】 [1] グループインデックス: 1～15 = SDフレーズ 16 = SVパスワードフレーズ
[2] コマンドポジション: 0～31 = コマンドポジションの指定

【解説】 指定したグループインデックスの、コマンドポジションに登録されたフレーズを削除します。フレーズ登録そのものが削除されますので、そのフレーズを学習させていた場合でも学習内容も削除されます。
フレーズを削除すると、もし削除したフレーズよりも大きい番号のコマンドポジション位置に別のフレーズが登録されている場合には、コマンドポジションが繰り上がります。よって、認識結果などもすべて繰り上がることになりますので注意が必要です。下記に例を示します。



5つのフレーズが登録されている　ポジション2のフレーズを削除　ポジション3以降のフレーズが繰り上がる

上記の例では、グループインデックス1に登録されている5つのフレーズのうち、コマンドポジション2のフレーズを削除した場合です。コマンドポジション2が削除されると、コマンドポジション3以降にあったフレーズのコマンドポジションすべて1つずつ繰り上がります。よって上記例では、例えば削除前はコマンドポジション3だった「OPEN」のフレーズが、削除後はコマンドポジション2になります。認識結果が変わりますので、ご注意ください。

【補足】 削除指定したコマンドポジションに登録されたフレーズがなかった場合には、"STS_INVALID"(76h)が返ります。

【期待戻り値】 STS_SUCCESS(6Fh) , "STS_INVALID"(76h)

【機能/動作】 *CMD_COUNT_SD* 指定したグループインデックス内にある登録フレーズ数を返します

【コマンド】 63h ('c')

【引数】 [1] グループインデックス: 1～15 = SDフレーズ 16 = SVパスワードフレーズ

【解説】 引数1で指定したグループインデックス内にある登録フレーズの個数を返します。
学習の有無にかかわらず登録されているフレーズの個数となります。

【補足】 戻り値は、"STS_COUNT"(63h)で返ります。最初に63hが返った後、"ARG_ACK"(20h)を送信後、登録されているフレーズの個数が返ります。

【期待戻り値】 "STS_COUNT"(63h)

【機能/動作】 *CMD_DUMP_SD* SDフレーズ・SVフレーズのラベルと学習回数を取得します

【コマンド】 70h ('p')

【引数】 [1] グループインデックス: 1～15 = SDフレーズ 16 = SVパスワードフレーズ
[2] コマンドポジション: 0～31 = コマンドポジションの指定

【解説】 指定したグループインデックスの、コマンドポジションに登録されたフレーズのラベルの長さとラベル、学習回数を所定の形式に従い返します。登録済みのフレーズのラベルとフレーズ学習の回数を知ることができます。

【補足】 戻り値は、"STS_DATA"(64h)で返ります。形式がありますので、詳しくは"STS_DATA"の項を参照ください。

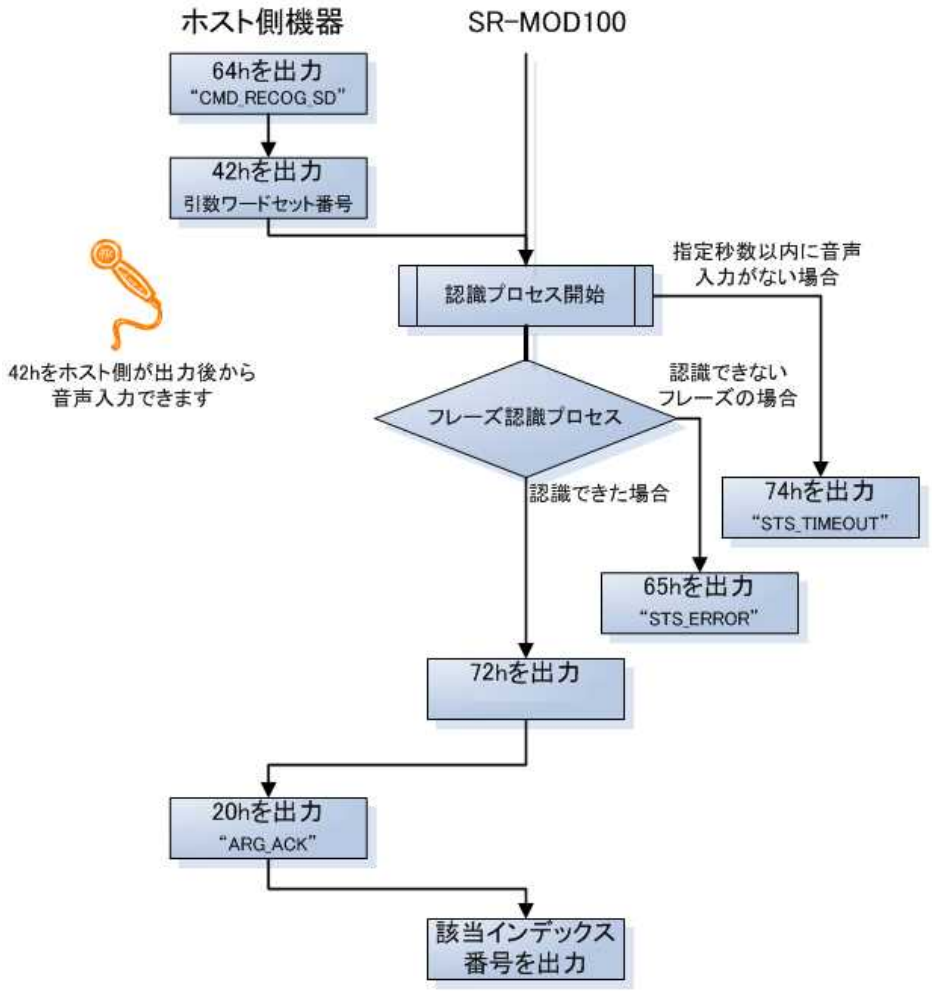
【期待戻り値】"STS_DATA"(64h)

【機能/動作】 *CMD_RECOG_SD* 特定話者認識プロセス及び話者特定認識プロセスを実行します

【コマンド】 64h ('d')

【引数】 1つ [1] グループインデックス: 1～15 = SDフレーズ 16 = SVパスワードフレーズ

【解説】 特定話者音声認識(SD)プロセス及び話者特定認識プロセス(SV)を実行します。
あらかじめ登録・学習させておいたグループインデックス内のフレーズについて、マイクから入力された音声と一致するかどうかの認識を実行します。引数には、グループインデックスを指定します。
音声認識プロセスは、引数の値を送信した時点から開始されます。タイムアップ時間を指定している場合には、引数の値送信後から音声入力がされる期間を時間カウントします。一連の動作フローを以下に記載します。
認識が正しく行われると、"STS_RESULT"(72h)が返り、ホスト側から"ARG_ACK"(20h)を送信すると、認識したフレーズに該当するインデックス値を返します。



上記例ではプレインストールフレーズとしてワードセット番号に1番(42h)を指定しています。例えばワードセット1番に「こんにちは」というフレーズが登録されていた場合、42h発行後に「こんにちは」と発声した場合、「こんにちは」というフレーズが登録されているリストのIndex番号に該当する値が返ります。

認識に成功し、リストにあるフレーズだった場合には"STS_RESULT"(72h)を、認識できないフレーズだった場合には"STS_ERROR"(65h)を、設定したタイムアウト時間以内に音声入力がなかった場合には、"STS_TIMEOUT"(74h)を出力します。
なお、注意が必要なのはフレーズの認識が出来た後、SR-MOD100は"STS_RESULT"(72h)を出力して認識できたことを通知してきますが、ホスト側が"ARG_ACK"(20h)を出力しないと、認識結果のインデックス番号を出力しないことにご注意ください。

【補足】 本コマンドを電源投入後に最初に実行する場合には、本コマンド実行の前に必ず"CMD_LANGUAGE"(6Ch)にて使用言語の設定を行っておく必要があります。使用言語の設定は、本来電源投入後電源が切断されなければ、1回だけでよいですが、本コマンド実行前に必ず言語設定を実行してから、本コマンドを実行するよう設計されても問題ありません。

【期待戻り値】STS_RESULT(72h) , STS_TIMEOUT(74h) , STS_ERROR(65h)

【機能/動作】 *CMD_QUERY_IO* 汎用I/Oピンの設定を行います

【コマンド】 71h ('q')

【引数】 [1] ピン番号: 1～3 = I/O1～I/O3ピンのピン番号を指定
[2] ピンモード: 0 = 出力Lレベル
1 = 出力Hレベル
2 = 入力ピン （ハイインピーダンス　プルアップ抵抗なし）
3 = 入力ピン （内部～10kΩプルアップ抵抗有効）
4 = 入力ピン （内部～200kΩプルアップ抵抗有効）

【解説】 汎用I/Oピンの状態を設定します。汎用I/Oピンの入力できる最大電圧は3.0Vです。5V系の回路とは直結できませんので、十分注意してください。
入力設定にした場合の状態把握については、"STS_PIN"(70h)で確認します。

【期待戻り値】 "STS_PIN"(70h)

■SR-MOD100からの戻り値一覧■

【内容名】 *STS_COUNT* カウントした結果の数量を返すコマンドです

【コマンド】 63h ('c')

【引数】 1つ [1] 整数0～31

【解説】 "CMD_COUNT_SD"(63h)等を実行した際、その結果としてカウントした値を返します。

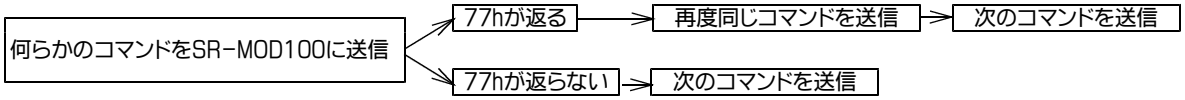
【関連コマンド】 CMD_COUNT_SD

【内容名】 *STS_AWAKEN* スリープモードからウェイクアップした時に返ります

【コマンド】 77h ('w')

【引数】 なし

【解説】 スリープモード(パワーダウンモード)から、SR-MOD100がウェイクアップした時に戻ります。
SR-MOD100を操作するホスト側機器は、必ずコマンドを送った際にこの77hが返らないかを確認するよう設計します。ホスト側からコマンドをSR-MOD100に送信した結果、77hが返った場合には、SR-MOD100はスリープモードであって、そのコマンドを受信したことでウェイクアップしたことを意味しています。その場合、SR-MOD100に送信したコマンドは破棄されていますので、再度同じコマンドを送信する必要があります。この仕組みがないと、コマンドが正しく実行されないことがありますので十分ご注意ください。
なお、SR-MOD100は、電源投入直後はスリープモードになっています。



【内容名】 *STS_DATA* 登録されたSD/SVフレーズの詳細情報を返します

【コマンド】 64h ('d')

【引数】 複数 [1] 音声認識学習情報: ※ −1=なし、 1−6=学習回数、 8=SD/SV競合フレーズ 16=SI競合フレーズ
[2] 競合コマンドのポジション: −1=なし、 0～31=競合しているフレーズのポジション
[3] ラベルの長さ 0～31 指定したインデックスに登録されているラベル(名前)の長さを返します
[4] ラベル名(ASCII) 指定したインデックスに登録されているラベル(名前)の文字列をASCIIで返します

【解説】 "CMD_DUMP_SD"(70h)を実行した際に、その結果を通知するコマンドです。主に登録されているフレーズの認識学習回数と、登録されてるラベル(名前)を返します。これによって、フレーズの学習が何回行われたか、また登録されているフレーズに付けられたラベルを取得することができます。

※引数1バイト目の学習情報で、値7(48h)及び0(41h)が返る場合があります。これは、−1(40h)と同様で、学習回数がないことを示しますが、より具体的には一度学習登録がされた後、その学習が削除されたときにその値が返ります。−1、0、7の値はいずれも学習が0回であることを示しています。

【関連コマンド】 CMD_DUMP_SD

【内容名】 *STS_ERROR* SI及びSDフレーズの認識に失敗した場合や学習に失敗したことを通知します

【コマンド】 65h ('e')

【引数】 複数 [1-2] エラーコード エラーコード8ビット分が上位4ビット、下位4ビットに分かれてそれぞれ1バイト、合計2バイトとして返ります(下記解説参照)

【解説】

"CMD_RECOG_SD"(64h)、"CMD_RECOG_SI"(69h)、"CMD_TRAIN_SD"(74h)などを実行した際に、その結果がエラーだった場合(例えば音声認識に失敗した場合など)に戻る戻り値です。65hに続いて、エラーコードが返ります。

エラーコードは特殊な形式で4ビットのコードが8ビット分(1バイト分)のデータとなって返ります。下記に詳細を記載します。
例えば、特定話者認識として"CMD_RECOG_SD"(64h)をホストからSR-MOD100に送信して、その後音声をマイクに発声します。
その結果、その発声した音声が、登録されていない音声だった場合、それは"ERR_RECOG_FAIL"(11h)となります。
この場合、SR-MOD100は、次のようなコードを返します。(次の値を取得するにはホスト側から20hを送信しているものとします。)

65h→42h→42h

42hはSR-MOD100の数値表現として"1"です。よって、最初の42hも1、次の42hも1なので0001 及び 0001 となります。
これが合計8ビットのエラーコードが上位と下位4ビットに分かれて返るという意味です。
これをまとめると、"00010001"となりますので、11hとなります。
例えば背景雑音の多い場所で複数の音声などが聞こえる場所で認識を行ってエラーになった場合には、"ERR_DATACOL_TOO_CHOPPY"(07h)が返ります。この場合には、65h→41h→48hとなります。これは、上位4ビットは0000、下位4ビットは0111となりますので、07hとなります。

【エラーコード】 エラーコードはエラーの内容を示す物ですが、あくまでも目安であり正確なエラー内容を示すものではありません。
エラーが発生したかどうかは、65hが返ったことで判別し、エラーコードは目安としてご利用ください。

03h	ERR_DATACOL_TOO_NOISY	雑音が多すぎます
04h	ERR_DATACOL_TOO_SOFT	入力音声が小さすぎます
05h	ERR_DATACOL_TOO_LOUD	入力音声が大きすぎます
06h	ERR_DATACOL_TOO_SOON	入力音声が速すぎます
07h	ERR_DATACOL_TOO_CHOPPY	要素が多すぎます、内容が複雑すぎます、認識が不安定です
11h	ERR_RECOG_FAIL	認識に失敗しました
12h	ERR_RECOG_LOW_CONF	認識結果が信頼できません
13h	ERR_RECOG_MID_CONF	認識結果が信頼できません
14h	ERR_RECOG_BAD_TEMPLATE	メモリーに記憶されたSD/SVコマンドが不正です
17h	ERR_RECOG_DURATION	間違ったパターン間隔です
80h	ERR_NOT_A_WORD	SIフレーズに登録されていない又はSIフレーズで認識できない音声です

主に03h～07hは音声信号取得に関する問題で発生します。使用場所周辺の騒音が大きい場合などで認識に問題がある場合に通知されます。11h～13hは音声レベルによるもので、このエラーが発生する場合には、本書最後部に記載の「音声認識の微調整」を実施してレベル設定をすると改善することがあります。
14h～17hは、内部メモリーの破損等によるものです。電源電圧の確認をした上で、一度すべての登録をし直す必要があります。
80hは、不特定話者認識(SI)を実行時に認識ができなかった場合に通知されるエラーです。登録されていないフレーズだったり、認識ができないような音声(音)だった場合に通知されます。

【関連ｺﾏﾝﾄﾞ】 "CMD_RECOG_SI" , "CMD_RECOG_SD" , "CMD_TRAIN_SD"

【内容名】 *STS_INVALID* 不正なコマンド又は値です

【コマンド】 76h ('v')

【引数】 なし

【解説】 コマンドとして定義されていない値や、範囲を逸脱した値等をSR-MOD100が受信した時に発行するコマンドです。
間違った値が送信されていますので、コマンドや送信している値について再度確認が必要です。

【内容名】 *STS_TIMEOUT* 実行した処理がタイムアウトした場合に発行されます

【コマンド】 74h ('t')

【引数】 なし

【解説】 音声認識開始コマンドや、音声認識フレーズ登録学習コマンドを発行後、一定時間何も音声入力がないなどし、タイムアップした時に戻る戻り値です。

【関連ｺﾏﾝﾄﾞ】 "CMD_RECOG_SI" , "CMD_RECOG_SD" , "CMD_TRAIN_SD"

【内容名】 *STS_INTERR* "CMD_BREAK"の結果として、処理を割込終了した時に戻ります

【コマンド】 69h ('i')

【引数】 なし

【解説】 音声認識開始コマンドや、音声認識フレーズ登録学習コマンドを発行後に"CMD_BREAK"(62h)をSR-MOD100が受信すると、直ちに現在の処理を強制的に終了して、本コマンドを返します。音声認識及びフレーズ学習が強制的に割込終了されたことを通知します。

【関連ｺﾏﾝﾄﾞ】 "CMD_BREAK"

【内容名】 *STS_SUCCESS* コマンドが正しく実行されたこと、及びOKを通知するコマンドです

【コマンド】 6Fh ('o')

【引数】 なし

【解説】 各種コマンド実行時に、そのコマンドの処理が問題無く完了したことを通知するコマンドです。コマンドが正しく受信され実行された場合や、コマンドに対する"OK"の場合に戻ります。

【内容名】 *STS_RESULT* SD及びSV認識での認識結果を通知するコマンドです。フレーズ学習時に既に似たフレーズが存在する場合にも返ります。

【コマンド】 72h ('r')

【引数】 1つ [1] コマンドポジション: 0～31 = コマンドポジション

【解説】 特定話者認識フレーズ(SD)並びに、話者特定認識(SV)において、その認識結果を通知するコマンドです。
認識が正常に行われ、結果が登録済みフレーズと一致すると、そのポジション位置を0～31の範囲で返します。
ホスト側から"CMD_RECOG_SD"を発行後、音声を入力し、その結果認識が正常にできるとこのコマンドをSR-MOD100は返します。認識に失敗した場合には、"STS_ERROR"(65h)を返します。
なお、72hが返った後、コマンドポジションの引数を取得する場合には、ホスト側から20hをSR-MOD100に送って取得します。

本コマンドにはもう1つ別の意味があります。特定話者認識フレーズ(SDフレーズ)を登録学習させた場合に、すでに同じグループ内に、よく似たフレーズが存在している場合、重複していることを示唆する目的で返ります。その場合、フレーズの登録及び学習結果の記憶、登録は可能ですが似たフレーズが既に存在しているため認識結果に間違いが生じる場合があります。

【関連ｺﾏﾝﾄﾞ】 "CMD_RECOG_SD" , "CMD_TRAIN_SD"

【内容名】 *STS_SIMILAR* SI認識での認識結果を通知するコマンドです。

【コマンド】 73h ('s')

【引数】 1つ [1] ワードインデックス: 0～31 = ワードインデックス

【解説】 不特定話者認識フレーズ(SI)において、その認識結果を通知するコマンドです。
認識が正常に行われ、結果が登録済みフレーズと一致又は一致していると考えられる似たフレーズと解析されると、そのワードインデックスの値を0～31の範囲で返します。
ホスト側から"CMD_RECOG_SI"を発行後、音声を入力し、その結果認識が正常にできるとこのコマンドをSR-MOD100は返します。認識に失敗した場合には、"STS_ERROR"(65h)を返します。
なお、72hが返った後、インデックス値の引数を取得する場合には、ホスト側から20hをSR-MOD100に送って取得します。

【関連ｺﾏﾝﾄﾞ】 "CMD_RECOG_SI" , "CMD_TRAIN_SD"

【内容名】 *STS_OUT_OF_MEM* 内蔵メモリーの容量がいっぱいです

【コマンド】 6Dh ('m')

【引数】 なし

【解説】 内蔵メモリーの容量がいっぱいになったことを通知するコマンドです。既に登録済みの認識フレーズを削除するなどの処置を行ってください。なお、登録数が少ないにもかかわらず本エラーが発生する場合には、内蔵メモリーの構造に問題が発生していることがありますので、一度"CMD_RESETALL"(72h)でメモリーを初期化してください。

※"CMD_RESETALL"を実行すると、登録されているすべてのデータは削除されます。

【関連ｺﾏﾝﾄﾞ】 "CMD_GROUP_SD"

【内容名】 *STS_PIN* 汎用I/Oピンを入力設定した際の入力状態を取得します

【コマンド】 70h ('p')

【引数】 1つ [1] ロジックレベル状態 0 = 入力状態Lowレベル、 1 = 入力状態Highレベル

【解説】 "CMD_QUERY_IO"(71h)で汎用I/Oピンの状態を入力設定にした際、I/Oピンのロジック状態がLowレベルか、Highレベルかの状態を通知するコマンドです。
"CMD_QUERY_IO"(71h)を発行した後、70hが返り、続いての引数を得るため20hを受信すると入力状態の結果を返します。

【関連ｺﾏﾝﾄﾞ】 "CMD_QUERY_IO"(71h)

使用例

本項では、使用例としてC言語でSR-MOD100を制御するプログラムを記述した例を紹介しています。

プログラム中の"SEND"はホスト側からSR-MOD100側にデータを送信していることを示します。また"RECEIVE"はSR-MOD100側から送信されたデータを受信していることを示します。"//"はコメントを示します。

(1)ウェイクアップ動作

```
//ウェイクアップするまで待機します

DO{
    SEND 'b';

}LOOP UNTIL REVEIVE = 'o'
```

コマンドo("STS_SUCCESS" 6Fh)を受信するまで"CMD_BREAK"を送信し続けます。多くの場合、すぐに"STS_SUCCESS"は受信しますが、SR-MOD100本体がハングアップしている場合にはそのままプログラムが停止してしまいますので、ループの試行回数を10回などに設定して10回送信しても"STS_SUCCESS"が戻らない場合には、再起動を促すなどの処置を取ると安全です。

なお電源投入直後は、スリープモードになっていますので、本コマンドを送信すると、最初の1回は必ず"STS_AWAKE"が戻りますので、その場合には再度"CMD_BREAK"を送信すると、続いて"STS_SUCCESS"を受信できます。

(2)電源起動後の使用言語等の設定

電源起動後は必ず使用言語の設定を行います。以下例は日本語に設定した場合です。またタイムアップ時間を5秒に設定しています。

```
//使用言語を設定します

SEND 'l';       //6Ch を送信
SEND 'C';       //日本語は 2 なので 43h(=C)を送信

IF RECEIVE = 'o' THEN OK ELSE ERROR;   //ACK の確認


//タイムアウトを 5 秒に設定します

SEND 'o';       //CMD_TIMEOUT 送信
SEND 'F';       //5 は 46h(=F)を送信

IF RECEIVE = 'o' THEN OK ELSE ERROR;   //ACK の確認
```

各種コマンド送信後は必ず、その後にSR-MOD100がACKである"STS_SUCCESS"(6Fh)を返すことを確認するよう設計してください。ACKが返らない場合には、そのコマンドの処理は正しく実行できていませんので、一度"CMD_BREAK"を送信して再度"STS_SUCCESS"が返ることを確認し、再度コマンドを送信するように設計します。戻り値は必ず確認する仕組みを作ってください。

(3)内蔵SIフレーズでの認識を実行する

ワードセット1にある不特定話者認識フレーズ(SI)の認識を実行してみましょう。タイムアップは5秒に設定されていますのでコマンド送信後5秒以内に発声しないと、タイムアップとなります。

```
//ワードセット 1 の SI フレーズ認識を開始

SEND 'i';       //SI 認識コマンド
SEND 'B';       //ワードセット 1

//ここで 5 秒以内に発声します
//例として"こんにちは"と発声した場合

result = RECEIVE;               //受信したコマンドを代入

IF result = 's' {               //認識成功
  SEND ' ';                     //スペース(20h)を送信
  command = RECEIVE - 'A';      //41h を引くと値になる

}ELSE IF result = 't'{          //タイムアウト

}ELSE IF result = 'e'{          //エラーを受信
  SEND ' ':                     //次の値を受信
  err = (RECEIVE - 'A') * 16;   //上位 4 ビット取得
  SEND ' ';
  err = (RECEIVE - 'A');        //下位 4 ビット取得

}ENDIF
```

コマンド受信後次のコマンドを受信する場合には、スペースの20hを送信します。

エラーコードは上位4ビットと下位4ビットに分かれて2バイトとして受信します。最初は上位4ビットが送られますので、受信した値から41hを引いた後、10h(=16)を乗算すると上位4ビットとなります。続いて、20hを送信して下位4ビットを取得します。例えばSIフレーズにないフレーズだった場合SR-MOD100は、エラーである"STS_ERROR"に続いて80hを2バイトに分けて、49h→41hと送信します。よって、

(49h-41h)×10h=80h　上位4ビット
(41h-41h)=0h　下位4ビット

となり、値は80hとなり、"ERR_NOT_A_WORD"になります。

(4)SDフレーズを新規に登録して、フレーズ学習させる

新規に特定話者認識フレーズ(SDフレーズ)を登録させて、そのフレーズに対して2回の反復学習をさせる場合のコード例を紹介します。

登録はグループ3、ポジション0に「DENGEN_ON」(電源オン)を追加し、2回学習させます。

```
//グループ 3 ポジション 0 に新規フレーズを追加

SEND 'g';
SEND 'D';       //グループ 3
SEND 'A';       //ポジション 0

//上記に対してラベルを割り当てます

SEND 'n';
SEND 'D';       //グループ 3
SEND 'A';       //ポジション 0
SEND 'J';       //文字数 9 文字 (4Ah)

SEND 'D';
SEND 'E';
SEND 'N';
SEND 'G';
SEND 'E';
SEND 'N';
SEND '_';
SEND 'O';
SEND 'N';       //ラベルここまで

IF RECEIVE = 'o' THEN OK ELSE ERROR
```

↓続きます

```
//登録した SD フレーズに対して学習を 2 回行います

SEND 't';
SEND 'D';
SEND 'A';

//ここで発声　「デンゲンオン」

result = RECEIVE

IF result = 'o'{                //問題無く完了
  OK

}ELSE IF result = 'r'{
  //認識は OK ですが似たよくフレーズが既に SD フレーズ
  //として登録されていると認識された場合は r が返る

  SEND ' ';     //20h 送信で競合するコマンドを取得
  command = RECEIVE - 'A';

}ELSE IF result = 't'{
  //タイムアウト

}ELSE IF result = 'e'{          //エラーを受信
  SEND ' ':                     //次の値を受信
  err = (RECEIVE - 'A') * 16; //上位 4 ビット取得
  SEND ' ';
  err = (RECEIVE - 'A');        //下位 4 ビット取得

}ENDIF
```

SDフレーズに対する学習は2回程度の反復学習をお奨めします。認識学習の結果は重要です。問題無く"STS_SUCCESS"が返ることを確認してください。"STS_RESULT"(r)が返る場合には、今学習させようとしたフレーズは既に同じグループ内に存在しているしている可能性があります。又はよく似たフレーズがある場合があり、その場合認識結果に間違いが生じたり、認識精度が悪くなることがありますので、フレーズの見直しが必要となります。

エラーの場合には、"STS_ERROR"の表を参照して、エラー対応をする処理を記述してください。

登録済みのSDフレーズの情報(学習回数、ラベル等)を取得したい場合には、"CMD_DUMP_SD"(70h)を使用します。

主な仕様

電源電圧:	DC5V(要安定化)
消費電流:	スリープモード時:約23μA
	通常動作待機時:約11mA
	認識動作時:約15mA
インターフェイス:	非同期式シリアル(TTLレベル)
対応OS:	WindowsXP/VISTA/7/8
生産国:	イタリア

使用上の注意

SR-MOD100の使用に際して、必ず下記の注意事項をお守りください。

①SR-MOD100を使用したことにより生じる、明示的及び潜在的な使危険性や、不確実性について、当方では予見することができません。使用する際には、お客様の責任においてこの製品を正しくお使いいただけますようお願い致します。

②当方では、SR-MOD100本体並びにサポートソフトウエア、本マニュアル等製品に関わるすべての製品、サービスについて安全性、正確性、有用性について一切の保証を行いません。

③SR-MOD100は、音声認識をすることを目的に開発、販売された製品ですが、その音声認識の信頼性は保証されていません。また使用については一般的な製品での使用の範囲に限定されます。本製品を宇宙、航空、原子力、燃焼制御、運輸、交通、各種安全装置、ライフサポート関連の医療機器等のように、特別な品質・信頼性を要求され、その故障や誤動作が直接人命を脅かしたり、人体に危害を及ぼす恐れのある用途向けには使用できません。

④SR-MOD100では様々な要因によっ音声認識の精度が悪くなったり、期待した性能を発揮できないことがあります。本製品を使用することによって生じた、もしくはこれに関連するいかなる直接・間接損害、懲罰的損害、その他データの破損や消失等を含むいかなる損害、損失についても、当方では一切責任を負いかねます。あらかじめご理解とご了承頂けますようお願い致します。

④本製品を使用した製品等を製造させる場合には、様々なフェイルセーフ機能(安全設計)を施して頂き、十分に機器のテストをした上で運用されますようお願い致します。また、データの損失や予期しない事態に備え、データのバックアップを行って頂けますようお願い致します。

サポート情報

よくお寄せいただくご質問については、当方のFAQページにまとめて掲載しております。下記のFAQページをご覧下さい。

http://www.microtechnica.tv/cgi-bin/faq.cgi

上記FAQにないご質問は以下のいずれかの方法でご質問をお寄せください。

■FAX番号　　03-3700-3548
■電子メール　　support@microtechnica.net

他社製品に関することや自作回路に関するご質問にはお答え致しかねますのであらかじめご了承ください。

なお、技術的なご質問をされる場合には必ずご使用環境や使用しているデバイスの型式、エラーが表示されている場合には正確なエラーメッセージをお知らせください。

マイクロテクニカ
〒158-0094　東京都世田谷区玉川1-3-10
TEL: 03-3700-3535　FAX: 03-3700-3548

(C)2015 Microtechnica All rights reserved